April 2004

The Impact of Calibration Error in Medical Decision Making

Final Report

Prepared for

National Institute of Standards and Technology
Chemical Science and Technology Laboratory
Building 101, Room A1000
Gaithersburg, MD 20899-0001

Prepared by

Michael P. Gallaher
Lee Rivers Mobley
RTI International
Health, Social, and Economics Research
Research Triangle Park, NC 27709

George G. Klee
Patricia Schryver
Mayo Clinic
Rochester, MN 55905

RTI Project Number 08236.001

RTI Project Number
08236.001

The Impact of Calibration Error in Medical Decision Making

Final Report

April 2004

Prepared for

National Institute of Standards and Technology
Chemical Science and Technology Laboratory
Building 101, Room A1000
Gaithersburg, MD 20899-0001

Prepared by

Michael P. Gallaher
Lee Rivers Mobley
RTI International*
Health, Social, and Economics Research
Research Triangle Park, NC 27709

George G. Klee
Patricia Schryver
Mayo Clinic
Rochester, MN 55905

*RTI International is a trade name of Research Triangle Institute.

Contents

Figures

Tables

GLOSSARY OF TERMS FOR LABORATORY TESTING (CLINICAL CHEMISTRY)

Accuracy is the measure of the variation in results from one method or one laboratory to another. Systematic inaccuracy is referred to as bias.

Bias is defined as the difference in means for two datasets resulting from systematic error.

Calibrators are solutions commonly supplied by equipment manufactures with know concentrations of the analyte. If in liquid form, the calibration solution will be comprise of the reference material, reagent, and solvent. If in solid form, the reference materials and reagent must be diluted with the solvent at the laboratory.

Calibration Curve plots absorbance at a specific wavelength against concentration of standards that have known concentrations. Calibration curves can be both linear and nonlinear.

Calibration Solution is the mixture of a primary standard (the solute) and a solvent.

Dilute Solution is a solution with relatively little solute.

Off-sets refers to a systematic shift in the mean of laboratory measurements. It is similar to bias in that it results from systematic error.

Precision is a measure of the random error associated with the test method and captures issues associated with reproducibility. Imprecision is typically expressed in terms of standard deviation or coefficient of variance.

Primary Reference Material (also referred to as primary standards) are highly purified materials that can be measured directly to produce a substance of exact know concentration.

Reference Material are substances that do not have the same level of purity of primary standards but each one has been characterized for certain chemical or physical properties and can be used in clinical chemistry.

Reagents are any chemical compound used as a reactant in a chemical reaction. Analytical reagents are those used in detecting, measuring, or analyzing other substances.

Uncertainty is also referred to as total error is a combination of both random and systematic error. One definition promoted by the EU community is 2 times the standard deviation plus the off-set (bias).

Executive Summary

Consensus guidelines and disease management strategies have standardized the medical approach to many common disorders. Unfortunately, the developers of medical guidelines have assumed that all laboratories function well and all test results are comparable. Medical guidelines seldom contain any information about the performance characteristics for key tests used in the diagnostic decision process. Guidelines typically have specific thresholds or "acceptable ranges," such as 8.9 to 10.1 mg/dL of calcium in tests to diagnose hypercalcemia, without any reference to measurement methodology or measurement standardization. Consequently, there is a false sense of security that the health care system assures adequate quality for laboratory tests.

Calibration error, leading to analytic bias, is a key parameter affecting the number of patients passing decision thresholds in practice guidelines. The Food and Drug Administration (FDA) requires that new tests perform equivalent to previously approved methods but does not require metrological traceability to reference methods. In addition, the Clinical Laboratory Improvement Amendments (CLIA) performance limits for proficiency tests are very wide, which allows large "between-lot" differences within methods and large "between-method" variations.

This study investigates the potential impact on health care costs from calibration error resulting in analytic bias in tests to measure serum calcium levels. Hypercalcemia is a medical condition caused by various disorders—most commonly hyperparathyroidism and cancer. The signs and symptoms of hypercalcemia are nonspecific; therefore, the clinician is very dependent on accurate laboratory measurements for detecting and evaluating this disorder. Medical guidelines recommend that

hypercalcemia be confirmed with follow-up procedures, such as intact parathyroid hormone (PTH) measurement, chest X-rays, 24-hour urinary calcium measurement, ionized calcium measurement, and thyroid imaging.

Based on analysis of over 89,000 patients receiving serum calcium tests at the Mayo Clinic in 1998–1999, we find that the number of follow-up procedures, and hence health care costs, is directly related to initial calcium test values. Based on interviews with laboratory managers and equipment manufactures, it was determined that calibration error has the potential to lead to bias of 0.1 to 0.5 mg/dL in up to 15 percent of calcium tests.

Analytic bias affects health care costs by increasing the number of follow-up tests performed for patients with elevated calcium levels. It is estimated that the cost impact associated with an analytical bias of 0.1 mg/dL could range from $8 to $31 per patient (receiving a calcium test). For an analytical bias of 0.5 mg/dL, which was the approximate upper bound identified during interviews, the potential health care cost increase ranged from $34 to $89 per patient having a calcium test.

With approximately 3.55 million patients per year receiving screening serum calcium tests being affected by systematic bias, the potential economic impacts range from $60 million to $199 million per year for analytic biases of 0.1 and 0.5 mg/dL, respectively.

1 Hypercalcemia

This study focuses on calibration errors in laboratory testing as they relate to the diagnosis of hypercalcemia. Calibration errors that positively skew calcium values in laboratory tests have the potential to significantly increase health care costs by increasing the number of follow-up procedures to diagnose hypercalcemia. Whereas depressed calcium levels can lead to conditions such as osteoporosis (weakening of the bones), hypercalciumia is much more prevalent in adults and is symptomatic of hyperparathyroidism, which, if untreated, can lead to kidney problems, bone fractures, and morbidity.

Section 1 begins with a description of the signs and symptoms of hypercalcemia and provides on overview of the typical follow-up procedures resulting from an elevated calcium test result. This description is followed by an overview of calcium testing procedures, focusing on the sources of error in the analytical phase.

1.1 HYPERCALCEMIA: SIGNS AND SYMPTOMS

Hypercalcemia is a condition that results in abnormally high levels of calcium in the blood (typically more than 10.2 mg per dL of blood). Although calcium plays an important role in developing and maintaining bones and other bodily functions, elevated levels of calcium have potentially harmful health implications. Normally, the body maintains a balance between the amount of calcium in food sources and the calcium already available in the body's tissues. This balance can be upset if the control systems regulating absorption, secretion, and bone resorption are malfunctioning because of disease.

Table 1-1 lists several signs and symptoms of hypercalcemia. Symptoms include extreme tiredness, mood swings, depression, confusion, nausea and vomiting, and increased urination. Elevated calcium levels can result in kidney stones, kidney damage, high blood pressure (secondary hypertension), and/or constricted arteries.

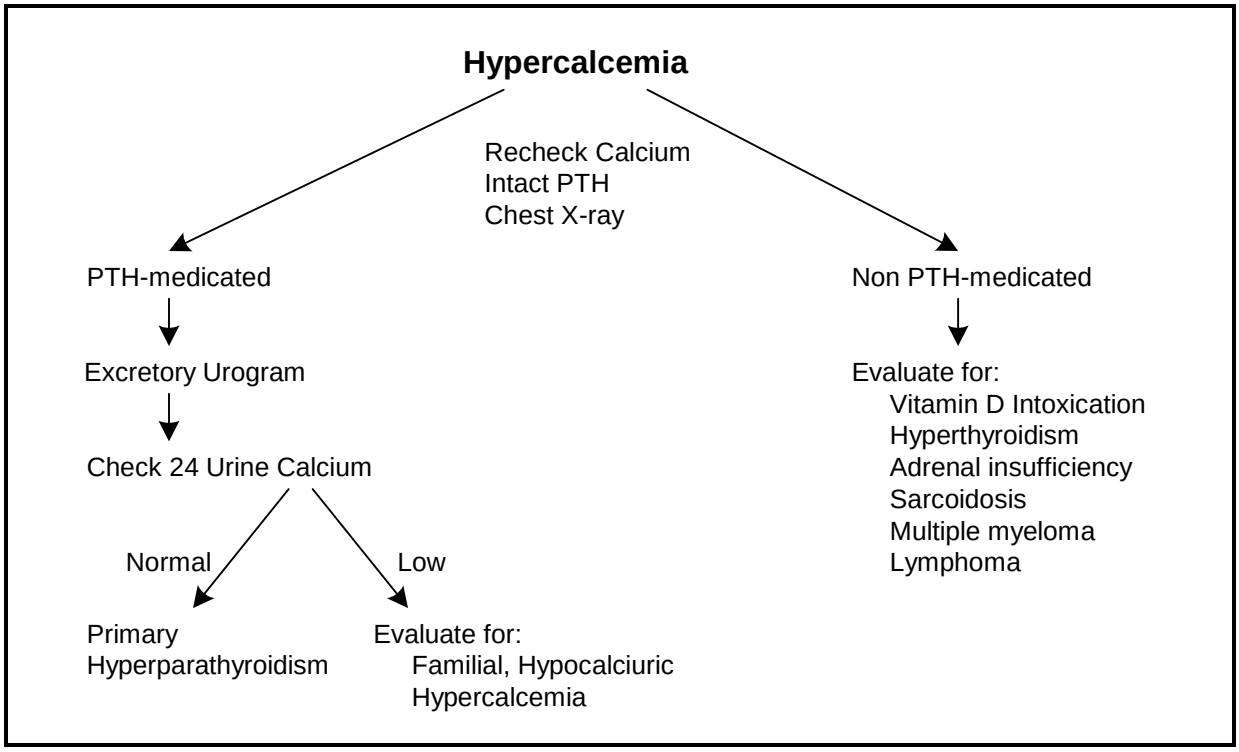
Table 1-1. Signs and Symptoms of Hypercalcemia

Mental	Neurologic and Skeletal	Gastrointestinal and Urological
Fatigue	Reduced muscle tone	Nausea
Obtundation	Muscle weakness	Vomiting
Apathy	Myalgia	Polyuria
Lethargy	Pain	Polydipsia
Confusion	Diminished deep tendon reflexes	Dehydration
Disorientation		Anorexia
Coma		Constipation

Hypercalcemia is also symptomatic of hyperparathyroidism, an endocrine disorder in which the parathyroid glands secrete too much parathyroid hormone (PTH). About 1 in every 2,000 adults has hyperparathyroidism, but in most cases, doctors do not know the cause of this disease. Frequent testing for hypercalcimia often occurs because untreated hyperparathyroidism can cause morbidity, and the early signs and symptoms of the disease are vague.

Because the risk factors for hyperparathyroidism are unknown, there is no way to prevent this disease, and hence frequent testing is common (HMS, 2001). Once an initial positive calcium test result emerges, a series of follow-up tests or procedures may be performed. These additional tests both reinforce the findings of the initial test and provide information to help diagnose the cause of the patient's elevated calcium levels. Figure 1-1 presents an example of typical follow-up procedures resulting from an elevated calcium test result. The first steps are to recheck the calcium level and conduct a PTH test and chest X-ray. If hyperparathyroidism is the diagnosis, medication protocols are initiated and further tests may be initiated. For example, if kidney stones are suspected, excretory urogram tests are conducted.

Figure 1-1. Example of Typical Follow-Up Procedures Resulting from an Elevated Calcium Test Result



Hand, feet, and skull X-rays may be ordered as follow-up tests to look for areas of diffuse bone demineralization, bone cysts, outer bone absorption, and erosion of the long bones of the fingers and toes. In addition to X-ray tests to document advanced hyperparathyroidism, the physician will probably order further tests to evaluate underlying complications.

Medically, the following procedures may be considered in resolving this differential diagnosis of hypercalcemia. The relative importance and the ordering sequences for these procedures depend on many circumstances, including patient presentation, patient physical examinations, and available facilities. Tests to be considered include the following:

1. repeat serum calcium
2. serum intact parathyroid hormones
3. chest x-ray
4. serum creatinine
5. excretory urogram
6. serum vitamin D level

7. thyroid stimulating hormone
8. free thryoxine
9. urine cortisol
10. serum angiotensin converting enzyme
11. serum protein electrophoresis
12. 24-hour urinary calcium
13. ultrasound of the neck

When blood calcium is only minimally elevated, the recommended treatment is to adopt a "wait and see" approach. In 1990, the National Institutes of Health convened a panel of experts that stated patients who are symptom-free, whose blood calcium is only slightly elevated, and whose kidneys and bones are normal may wish to talk to their doctor about long-term monitoring. Monitoring consists of clinical evaluation and measurement of calcium levels and kidney function every 1 to 2 years. If the disease shows no signs of worsening after 1 to 3 years of monitoring, the interval between exams may be lengthened. If the patient and doctor choose long-term monitoring, the patient should try to drink lots of water, get plenty of exercise, and avoid certain diuretics, such as thiazides.

For higher calcium levels (and/or if indicated by other tests), surgery is recommended to remove the enlarged gland(s). Surgery cures hyperparathyroidism in 95 percent of cases and has a low complication rate when performed by surgeons experienced with this condition. About 1 percent of patients undergoing this surgery experience damage to the nerves controlling the vocal cords, which can affect speech. One to 5 percent of patients who have the surgery develop chronic low calcium levels, which may require treatment with calcium and/or vitamin D.

Although a benign parathyroid tumor is 85 times more likely than a malignant one, in rare cases (1 to 2 percent of adults with hyperparathyroidism), pathologist's review of the removed tissue indicates cancer. This form of cancer usually strikes adults in their 40s and 50s and can spread quickly to other areas of the body, resulting in death. The survival rate is about 60 percent if detected within 5 years and drops to about 40 percent if detected within 10 years.

1.2 TESTING FOR HYPERCALCEMIA

Patients typically provide a blood or urine sample to be tested for hypercalcemia. Calcium in blood serum is found in three different forms: ionized, complexed, and protein bound. Most tests determine the

concentration of total calcium, which is the sum of the three forms and is a measure of total serum calcium.

Reference ranges are used to determine how much calcium is expected and natural within the specimen. Levels outside the reference ranges usually indicate a need for further testing. For adults, a typical calcium reference range is 8.9 mg/dl to 10.1 mg/dl. Critical action levels for calcium occur if the calcium in blood is below 7.0 mg/dl or rises above 13.0 mg/dl.

The testing process can be segmented into three phases: pre-analytic, analytic, and post-analytic. (Table 1-2 summarizes the testing procedure activities.) Uncertainty (which includes both random and systematic error) is primarily introduced in the pre-analytic and analytic phases. The focus of this study is on systematic errors introduced in the analytic phase of testing. This is also referred to as "calibration error." Calibration error introduces an analytic bias into laboratory test results potentially leading to an increased number of false positives for hypercalcemia, and hence increasing health care costs.

In the ***pre-analytic stage***, the specimen is collected from the patient and then stored in a holding location to await the analytic phase. Errors introduced in this phase are typically a function of human error rather than a function of testing techniques. Examples of errors include inappropriate collection techniques, incorrect labeling, poor storage techniques, or cross-contamination. Improved testing techniques and reference models in the analytic phase are not expected to influence the collection procedures and/or other phases of the pre-analytic process. Even though this study does not focus on errors generated during the pre-analytic phase, it should be noted that errors occurring at this stage can influence final test results and generate significant costs due to retesting and inaccurate diagnoses.[1]

The ***analytic phase*** begins once the sample has been collected and has reached the testing facility. The first step of the analytic phase involves mixing the specimen and reference materials with accurate amounts of the appropriate reagents for testing. For automated calcium measurements, reagents usually consist of o-cresolphtalein complexone *(*o-CPC), nonreactive surfactant (CAPS), and 8-hydroxyquinoline. At this point, both the specimen and the reference materials are tested

[1]It should also be noted that uncertainty may exist at the patient level based on patient-specific characteristics; however, this is also excluded from the scope of this study.

Table 1-2. Testing Procedure Activities

	Pre-Analytic Phase	Analytic Phase	Post-Analytic Phase
Diagnostic Reagent/ Equipment Companies		1. Design specification for diagnostic instruments and reagents 2. Quality control of instruments and reagents 3. Quality control of reagents by lot 4. Assignment of values to calibrate by lot 5. Respond to service requests for equipment and reagents	1. Monitor problems from field to plan design improvements
Patients/ Processing	1. Preparation – Fasting – Stabilizing – Provocative stimulation 2. Collection of specimen 3. Processing and storage of specimen		
Laboratory		1. Validation of instrument (daily) 2. Validation of reagent (daily) 3. Calibration of instrument (daily) 4. Quality control of instrument (every 100 tests). Loading of reagent and controls 5. Mixing and pipetting of specimen 6. Analysis 7. Verification/release of sample results	1. Archive sample 2. Record quality control 3. Monitor quality control and distribution of test results

separately. The next step in the process is typically to measure spectrophotometric signals and compare the results from the sample to the results generated by using the reference model. The final result is the level of calcium contained in the specimen.

Conceptually, most manufacturers' instruments use calibration curves for determining the calcium concentration in the sample. Identifying calibration curves depends on both the absorbance of the sample and calibration solutions, and the concentration of the calibration solutions. A linear, two-point calibration curve can be defined by the following equation:

$$C_x = C_0 + \left[\frac{A_s - A_0}{A_{cal} - A_0}\right] * (C_{cal} - C_0) \qquad (1\text{-}1)$$

where

C_x = total concentration of calcium in the sample solution,

C_0 = total concentration of calcium in the solution used to establish the zero-point of the calibration curve,

A_S = normalized and blank-corrected absorbance signal of sample solution,

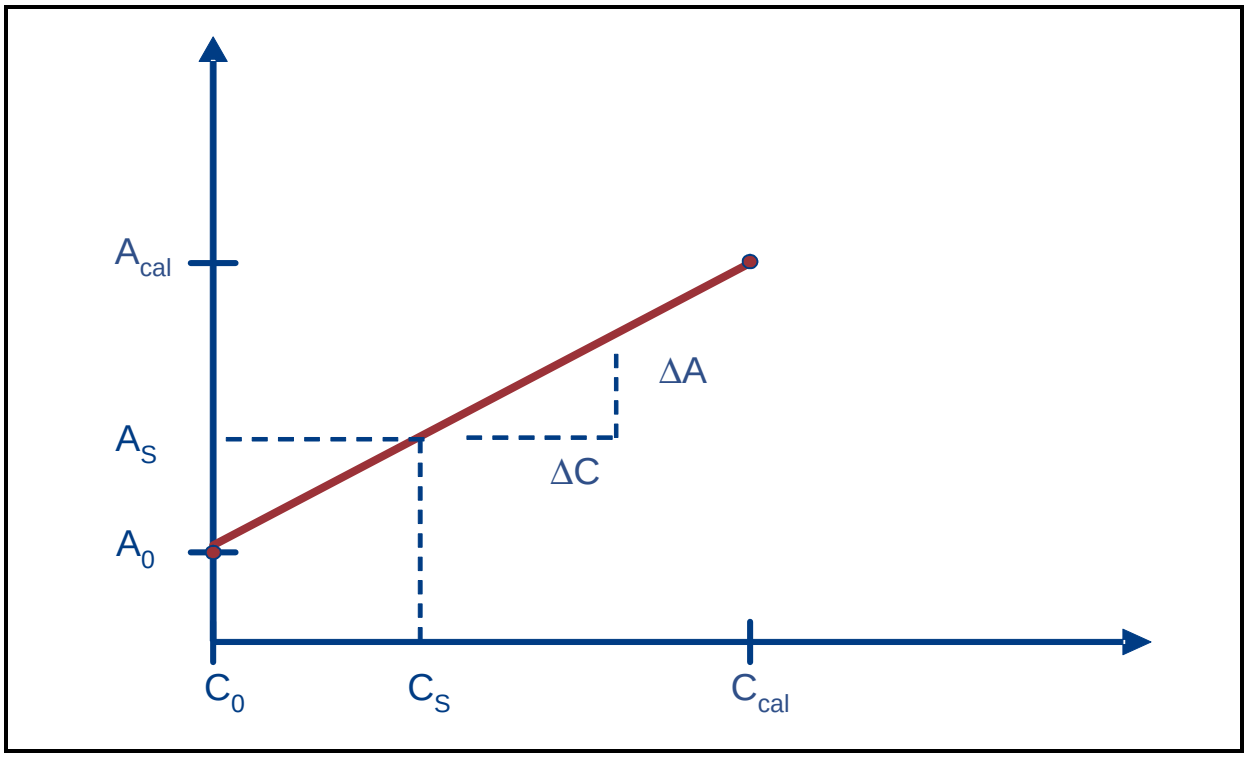
A_0 = absorbance signal from reagents,

A_{cal} = normalized and blank-corrected absorbance signal of calibrator solution, and

C_{cal} = total concentration of calcium in the calibrator.

The relationship in Eq. (1.1) is also represented graphically in Figure 1-2.

Figure 1-2. Illustration of a Two-Point Calibration Curve



Systematic error is typically associated with the calibration reference materials and reagents. Sources of systematic error are further discussed in Section 1.2.1. Random error is associated with measurement of the absorbance signals and largely depends on the

reference method used (photometry, atomic absorbtion spectroscopy, etc.). An overview of reference methods commonly used for calcium tests is presented in Appendix A.

The ***post-analytic phase*** involves distributing and archiving the test results. In this phase, inadequate information systems can lead to unnecessary duplication or "redundant" lab tests when the clinician ordering the tests is not aware of the orders or prior test results from other clinicians (Bates et al., 1998; Van Loon et al., 1999). However, the potential benefits of electronic information systems to link patient records are not a major focus of this study.

1.2.1 Sources of Systematic Error in the Analytic Phase

Systematic error in the testing process is primarily associated with calibration activities and materials. Sources of systematic error introduced in the analytic phase include

- calibrators (lot-to-lot variation),
- traceability of reference materials,
- measurement reagents,
- matrix effects, and
- changes in instrument calibration (drift).

These factors are discussed below.

Calibrators

Calibrators link the absorbance measure of the testing equipment with known concentrations and are used to develop calibration curves. The absorbance measure for the calibration is dependent on the contents of the solution, which typically contains the reference material (known value of calcium), reagent, and solvent. Manufacturers of testing instruments usually provide the calibrator solutions, or solutes that are mixed with diluted solvent at the laboratory. These measurement standards from the manufacturer are accompanied by a certificate with information about the values of the calcium concentration in the calibrators.

Calibrators are produced in large batches (referred to as "lots") and then segmented into individual parcels for periodic use over time. Lot-to-lot variations in the calcium and reagent concentrations of calibrators can lead to systematic measurement error (bias) over the lifetime of individual lots.

In addition, the absorbance reading of the dilutant solution is used to establish the zero-point (baseline) of the calibration curve. The difference between the absorbance reading of the calibrator, specimen, and the baseline determines the test result; thus, the test result can be sensitive to concentration errors in either the calibrator or dilutant.

Traceability of Reference Materials

Traceability establishes a link between secondary reference materials and the primary standard. NIST has developed certified standard reference materials (SRMs) for use in clinical chemistry laboratories, including a human serum standard reference material (SRM 909b1 and 909b2) for calcium. The use of standard reference materials is an important part of quality assurance programs that support the verification of the accuracy of specific measurements.

Measurement Reagents

Reagents are also commonly supplied by the equipment manufacturers and mixed with the sample prior to testing. Because the reagent also influences the absorbance reading, variations in reagent concentration or volume across batches can lead to systematic error in test results.

Matrix Effects

It is impossible to have calibrators with exactly the same properties as the patient sample. Even if the concentration of calcium in the sample and the reference material were the same, the concentration of other naturally occurring components may be different. Any variation in composition between the sample and reference material can result in a difference in instrument response and hence an error. These differences are referred to as “matrix effects.”

In addition to the simple heterogeneity of patient properties, some analytical processes can generate systematic matrix effects. A commonly cited practice leading to matrix effects is the process of freeze-drying and reconstituting calibrators that can lead to changes in composition.

Changes in Instrument Calibration (Drift)

Instruments are typically calibrated on a weekly or monthly basis, and control samples are typically measured every 6 hours. The information from the control samples is logged into a chart as part of the internal quality control procedure. Changes in the instrument readings (“drift”) between calibrations constitute a source of uncertainty. If measurement

on a control sample indicates that the calibrator's set-point has changed and falls outside an accepted interval, additional procedures are performed and recalibration may be needed.

2 Methodology for Estimating Impacts

The primary objective of this study is to investigate the economic impact of calibration error associated with laboratory tests of calcium levels. Calibration error includes both random (variance) and systematic (bias) error. However, systematic error has the most significant impact on medical decision making because it leads to analytic bias that shifts all test values and can cause more patient results to be beyond the clinical decision limit.[1]

To estimate the impact of systematic error on health care costs, cost functions are developed in Section 3 that express the average expenditures on follow-up procedures as a function of the initial calcium test value. The general concept is that when patients receive a calcium test value outside the reference range, physicians are likely to order a follow-up test. Thus, elevated calcium values have the effect of increasing the likelihood of additional follow-up tests. However, because many other symptoms are also considered in the diagnosis, there is no discreet threshold where specific actions are triggered. This relationship is illustrated in Figure 2-1, where the probability of follow-up tests are an increasing function of the initial calcium test value.

In addition to elevated calcium values increasing the probability of follow-up activities, elevated calcium levels are also likely to increase the number and complexity of follow-up activities. For example, a patient with a slightly elevated level of calcium may simply be retested, whereas a patient with a significantly elevated calcium level may have multiple follow-up activities, including PTH tests or chest X-rays. Incorporating these factors yields a health care cost function, where expected follow up

[1] Petersen et al. (1997) found that analytical bias has a significant impact on diagnostic performance and can lead to an unacceptable percentage of diagnostic misclassifications (false positives and false negatives) based on current standardization methods and quality specifications.

health care costs are an increasing function of initial calcium test values (see Figure 2-2).

Figure 2-1. Probability of Follow-Up Tests as a Function of Initial Calcium Level

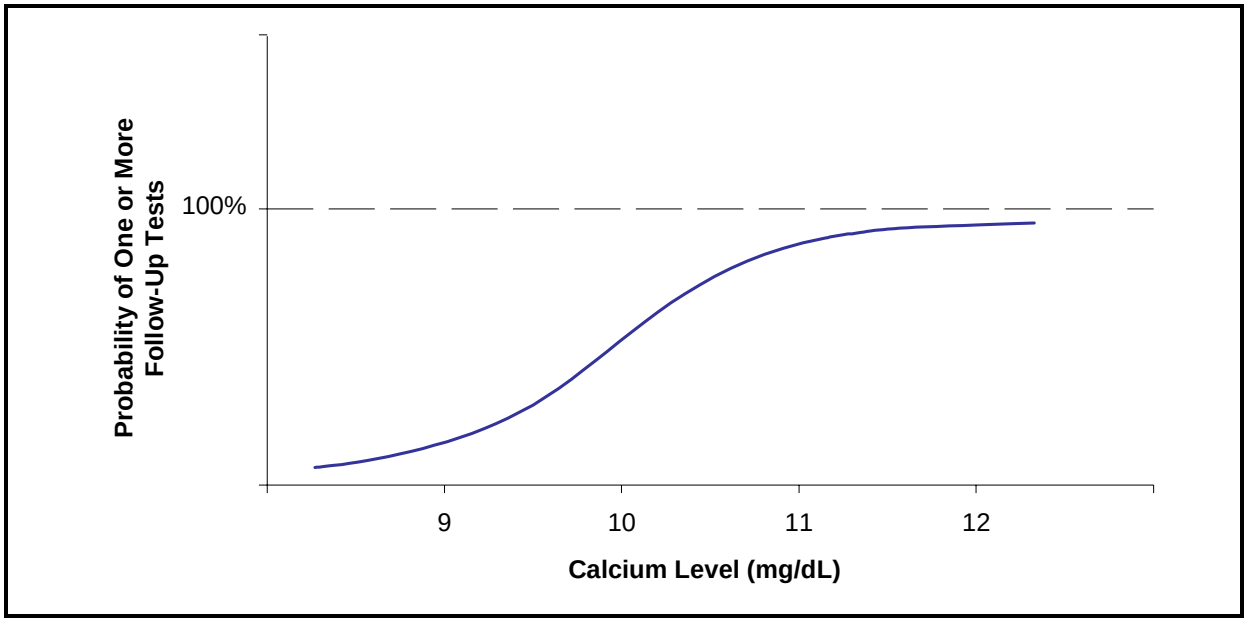
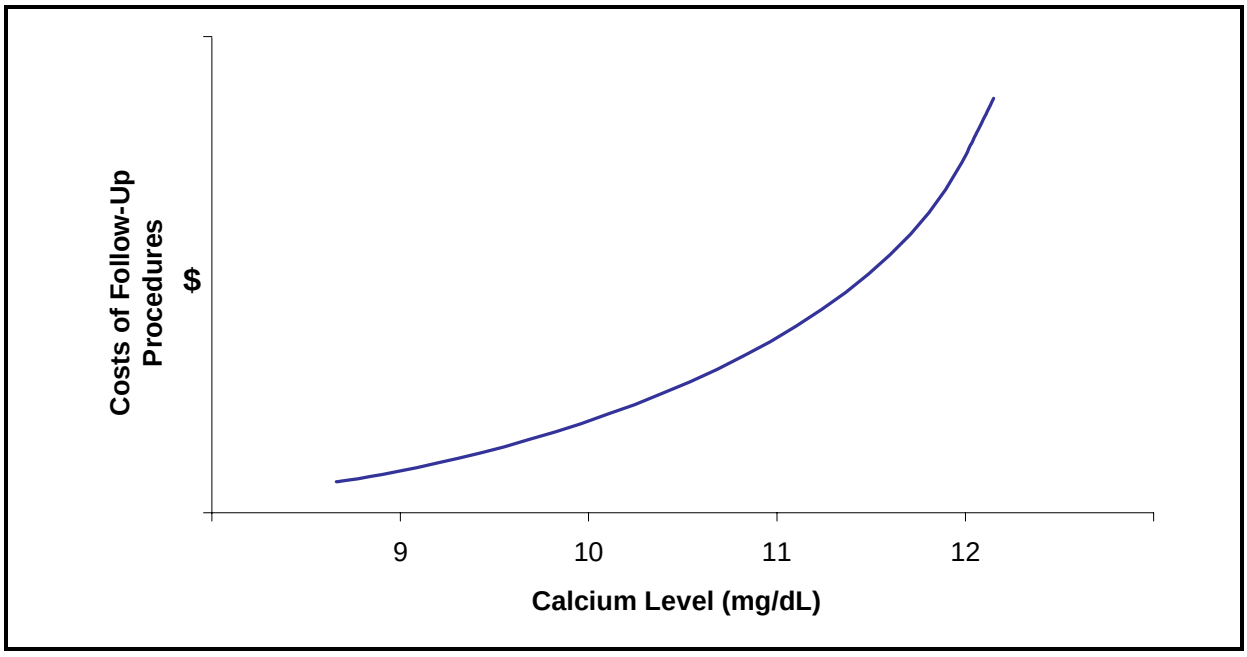


Figure 2-2. Costs of Follow-Up Test as a Function of Initial Calcium Level



The impact of a systematic calibration error can be illustrated as an upward shift in the follow-up health care cost function. As shown in Figure 2-3, a positive systematic error (also known as an "offset") will shift the cost function, increasing expected health care costs. For example, a patient with an actual calcium level of 11.0 mg/dl will now receive follow-up tests associated with a calcium value of 11.5 mg/dl. This can be translated into an incremental expected cost function as shown in Figure 2-4. The incremental cost function can then be used in conjunction with the frequency distribution of calcium levels in the U.S. population to calculate the incremental health care costs associated with systematic calibration error. Incremental health care costs can be estimated for any potential level of calibration error using this approach once the cost function has been developed.

Section 3 presents the data and analysis steps used to develop the cost function. Section 4 discusses the systematic error ranges used in the analysis that were developed from interviews with industry experts.

Figure 2-3. Shift in the Follow-Up Cost Function due to Analytic Bias

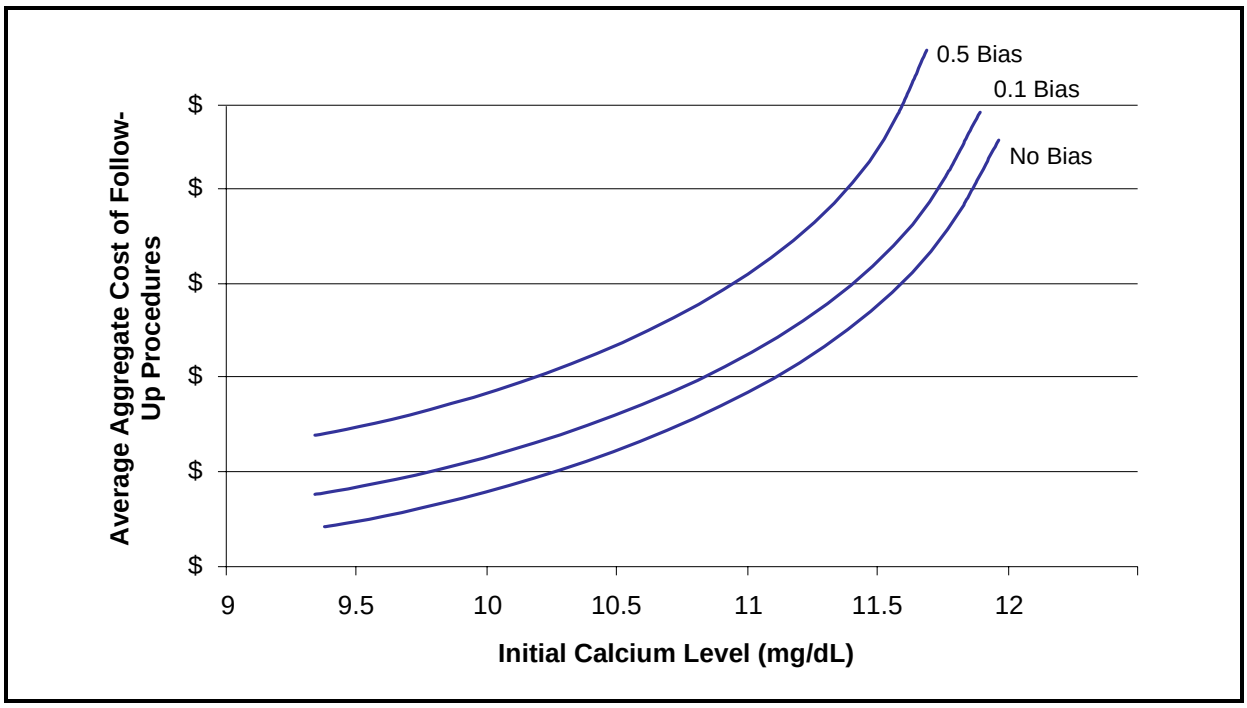
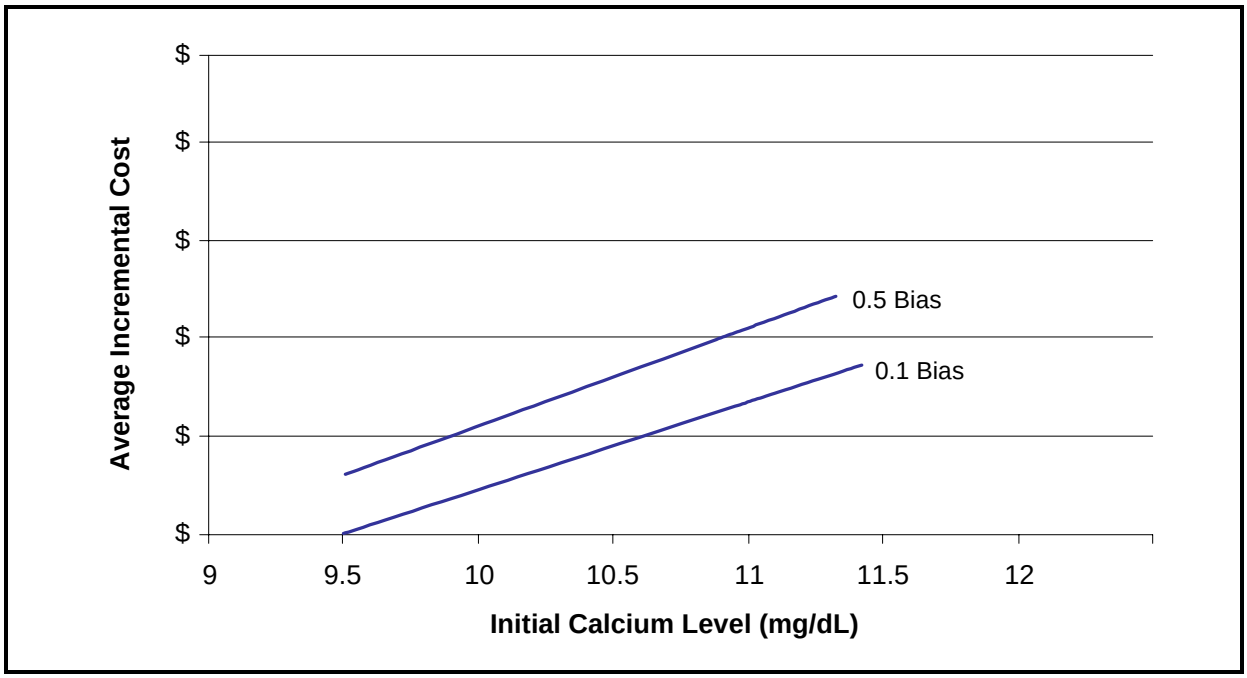


Figure 2-4. Incremental Cost per Patient Associated with 0.1 and 0.5 mg/dL Biases



3 Development of Cost Function

A combination of medical decision logic and data-driven statistical associations were used to model the relationship between initial calcium test values and health care costs. Cost functions were developed for four population subgroups: female-Medicare, male-Medicare, female-private insurance, and male-private insurance. Males and females were modeled separately because of differences in calcium value distributions and differences in the profiles for follow-up tests. Medicare and private insurance patients were partitioned because of costs (charging) differences.

The following steps were used:

1. Identify patients with initial calcium tests received in 1998 and 1999 using data from Mayo Clinic's patient population. Compile the population distributions in each of the four subgroups as a function of the initial serum calcium concentration test result.
2. Identify follow-up tests and procedures that were ordered more frequently in patients with hypercalcemia. Establish frequency response curves relating the number of follow-up procedures ordered as a function of the initial calcium level for the four subgroups of patients.
3. Assign Medicare and private insurance reimbursement rates to follow-up procedures, and use these to calculate total procedure costs for each group of patients. The Medicare reimbursement rates were obtained from the national Current Procedural Terminology (CPT4) fee schedule. The private insurance reimbursement rates were calculated from a weighted average of private payer reimbursement rates from nine geographic regions.
4. Analyze the impact of case severity on the cost model by looking at the effect of the number of replicate tests per patient. Then use these data to identify a cutoff value for the maximum number of tests per procedure to be included in the cost analysis. This will eliminate tests not associated with the initial calcium value.

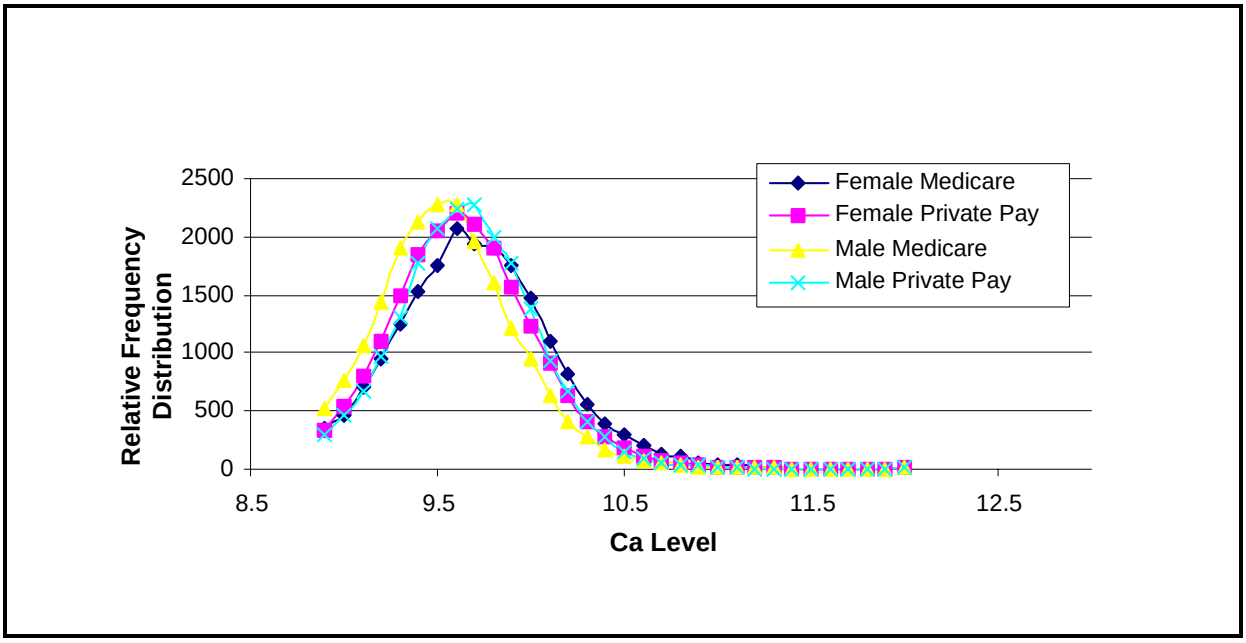
The cost function was then used to simulate the effect of systematic error of calcium measurements on health care costs.

3.1 FREQUENCY DISTRIBUTION OF INITIAL CALCIUM VALUES FOR EACH SUBGROUP

The analysis population was developed from electronic laboratory and billing records of patients seen at the Mayo Clinic in 1998 and 1999. A total of 89,083 adult patients 18 years and older were identified who had at least one serum calcium test with a value greater than or equal to 8.9 mg/dL performed during these 2 years and who had given research authorization. The calcium values, test dates, age, and gender were extracted from the laboratory file, along with follow-up tests and procedures (specified as CPT4 codes) for the 12 months following the initial calcium test.

Patients were grouped into 0.1 mg/dL cells to form a calcium value frequency distribution. The calcium values are reported to one decimal place in units of mg/dL. Figure 3-1 shows the frequency distributions of initial calcium values for each of the payment-gender subgroups (female-Medicare, male-Medicare, female-private insurance, male-private insurance). The normal reference range for calcium is 8.9 to 10.1 mg/dL.

Figure 3-1. Relative Frequency Distribution of Calcium Values



The majority of the patients in each subgroup had calcium measurements in the normal range of 8.9 to 10.1 mg/dL. Approximately 10 percent of patients had calcium measurements above 10.1 mg/dL and only 0.2 percent of the patients had calcium values above 11.3 mg/dL.

3.2 FOLLOW-UP TESTS AND PROCEDURES ORDERED MORE FREQUENTLY IN PATIENTS WITH HYPERCALCEMIA

For each patient with an initial calcium test in 1998 or 1999, follow-up tests and procedures were extracted from the laboratory file and matched with their CPT4 codes found in the billing file. All tests and procedures occurring during the 12 months following the initial calcium test were extracted. On average, each patient had 81 CPT4 procedures within the following 12 months. The median number of procedures per patient was 62, with the top 10 percent of patients accounted for 42 percent of the procedures. However, as discussed below, out of the hundreds of different CPT4 procedures present in the patient database, only 26 were determined to be correlated with hypercalcemia and were used in the cost analysis.

3.2.1 Removing Procedures Not Positively Correlated with Hypercalcemia

For each procedure, the relative frequencies were cross-plotted against the initial calcium values and linear regression slopes were calculated. Twenty-six CPT4 procedures with positive slopes greater than 0.010 were identified as empirically associated with hypercalcemia. These procedures, shown in Table 3-1 (along with the regression slope), were used in the health care cost analysis. The remaining procedures were dropped from the analysis.

Many of the tests and procedures clinically associated with hypercalcemia are also ordered for numerous other medical conditions. This dilutes the association with hypercalcemia so that the test ordering was not statistically correlated with higher calcium values. For example, thyroid function tests are used to evaluate patients with hypercalcemia, but they are also commonly conducted to evaluate numerous other medical problems such as fatigue and eye problems. Causality is an important concern of this study because the objective is to identify additional tests resulting from elevated calcium test results. However, if most of the tests are ordered as a result of other medical conditions, the

Table 3-1. Procedures Positively Correlated with Hypercalcemia

	Procedure	CPT4 Code	Regression Slope[a]
1	Explore Parathyroid Glands	60505	0.221
2	Chest X-Ray	71020	0.054
3	Nuclear Scan of Parathyroid	78070	0.101
4	Assay Serum Albumin	82040	0.076
5	Angiotensin Enzyme Test	82164	0.033
6	Assay Calcium in Blood	82310	1.919
7	Assay Calcium in Urine	82340	0.219
8	Assay Blood Carbon Dioxide	82374	0.082
9	Assay Blood Chlorides	82435	0.073
10	Assay Cpk in Blood	82550	0.067
11	Assay Blood Creatinine	82565	1.053
12	Assay Urine Creatinine	82570	0.114
13	Assay Ferritin	82728	0.022
14	Glucose, Blood,-Gluc. Monitoring Dev.	82962	0.057
15	Assay Blood Magnesium	83735	0.164
16	Assay of Parathormone (PTH)	83970	0.331
17	Assay Alkaline Phosphatase	84075	0.487
18	Assay Blood Phosphorus	84100	0.282
19	Assay Blood Potassium	84132	1.946
20	Assay Blood Sodium	84295	0.652
21	Assay Bun	84520	0.060
22	Automated Hemogram	85025	1.932
23	Prothrombin Time	85610	0.107
24	Blood Typing; Abo	86900	0.082
25	Culture Specimen, Bacteria	87070	0.038
26	Urine Culture, Colony Count	87086	0.063

[a]Slope = Linear regression for the cross-plot slope of the ratio of the number of the patients having that CPT4 code divided by the number of patients having that calcium (Y), versus the initial serum calcium level (X).

impact of measurement bias in calcium tests on health care costs is unclear. These other tests and procedures were screened out as described below.

Conversely, several tests were found to have strong statistical associations with elevated calcium test results that were not initially hypothesized to be linked to hypercalcemia. The explanation for these test orders is not known, but some may be due to clusters of ordering

patterns and/or the coexistence of other diseases in patients with hypercalcemia.

3.2.2 Follow-Up Procedures as a Function of Initial Calcium Value

For each calcium value, the total number of patients and the number of patients having each of the possible procedures (as defined by unique CPT4 codes) were enumerated. The relative ordering frequency of these procedure codes (e.g., number of procedures divided by number of patients having that value of initial calcium) was calculated for each of the 24 initial calcium value intervals. For example, Table 3-2 shows that the number of PTH tests per patient increases as the initial calcium value increases. This finding is similar to the probability of receiving a parathormone test, given the initial calcium value.

Table 3-2. Number of PTH Follow-Up Tests per Patient

Calcium Value (mg/dL)	Number of Patients	Number of PTH Procedures	Number of Procedures per Patient
8.9-10	78,232	1,398	0.018
10.1	4,070	112	0.028
10.2	2,862	306	0.107
10.3	1,909	315	0.165
10.4	1,282	360	0.281
10.5	850	277	0.326
10.6	568	224	0.394
10.7	351	165	0.470
10.8	258	126	0.488
10.9	179	98	0.547
11	133	77	0.579
11.1	94	43	0.457
11.2	66	39	0.591
11.3	54	38	0.704
11.4	30	18	0.600
11.5	29	19	0.655
11.6	29	20	0.690
11.7	17	10	0.588
11.8	14	10	0.714
11.9	32	25	0.781
≥12	85	52	0.612

3.2.3 Subgroup-Specific Procedure Frequency Functions

After removing procedures not positively correlated with hypercalcemia, separate frequency functions were then developed for the four subgroups (male-Medicare, female-Medicare, male-private insurance, and female-private insurance) depicting the relative frequency of receiving each follow-up procedure as a function of initial calcium concentration. One of the reasons for developing these functions was to examine whether different subgroups had similar follow-up procedures as a function of initial calcium values.

The frequency functions were developed using a strategy similar to that described above, with a calcium value range of 8.9 to 12.0 mg/dL. For each of these 26 discrete calcium values, the ratio of the number of procedures ordered (in the follow-up 12 months) to the total number of patients having that initial calcium value was derived. Least squares techniques were used to fit curves to each of these ratios. Figure 3-2 provides a representative example of these functions for the parathyroid hormone assay. Separate functions are shown for each of the four gender-payment subgroups. In general, the subgroups had similar frequency functions; this was the trend for all of the 26 follow-up procedures included in the cost model.

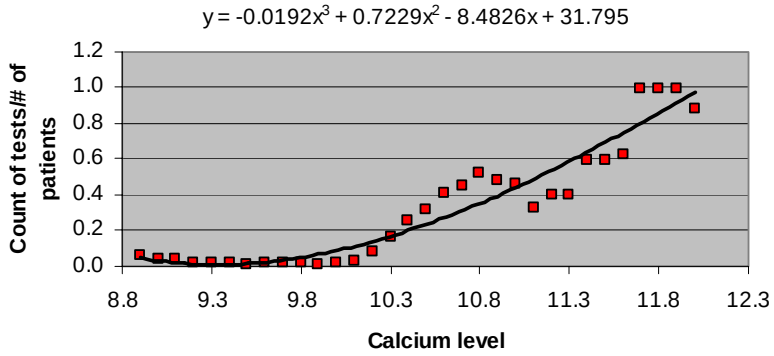
The slopes of the curves are important because they determine the incremental costs associated with bias or uncertainty. For example, if the curves are flat, as they typically are around the normal range, this implies that the same number of follow-up tests is ordered regardless of where the test results fall within this local region. Thus, within this flat range, systematic error or bias has minimal to no impact on follow-up test costs. In contrast, as the slope becomes steep, as in the elevated calcium regions, bias (e.g., shifting a measurement from 11.3 to 11.4, for example) will increase the number of follow-up tests and lead to higher costs.

3.3 ASSIGNMENT OF MEDICARE AND PRIVATE INSURANCE COSTS TO CPT4 PROCEDURES

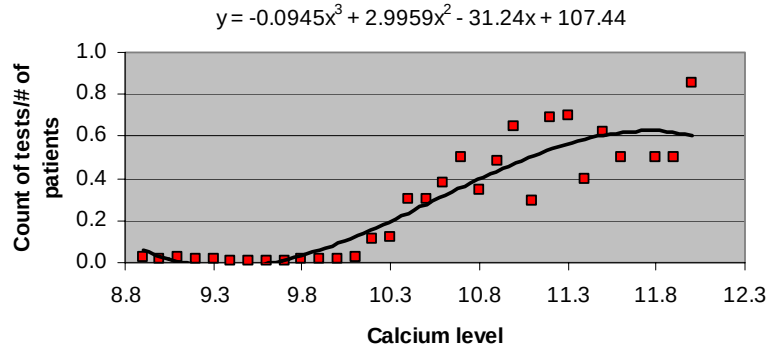
Table 3-3 shows the national Medicare fees and the assigned private insurance costs for the CPT4 codes used in this model. The source of the cost data is the 2000 Medicare Clinical Diagnostic Laboratory Fee Schedule. The fee schedule was obtained from the Centers for Medicare & Medicaid Services (CMS) website (http://cms.hhs.gov/providers/pufdownload/default.asp#labfee). The Medicare fee schedule

Figure 3-2. Parathyroid Hormone Assay

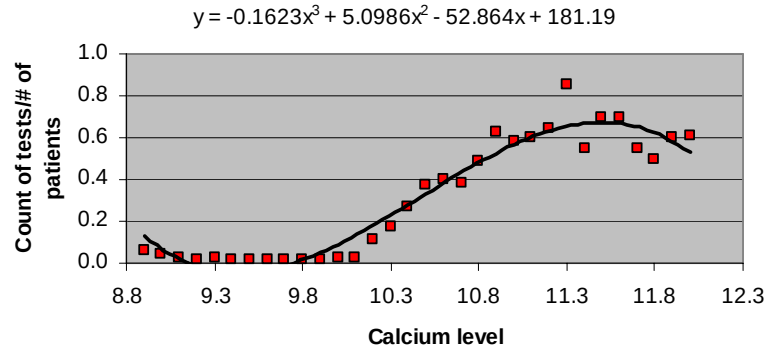
a) Males—Private Insurance



b) Males—Medicare



c) Females—Private Insurance



d) Females—Medicare

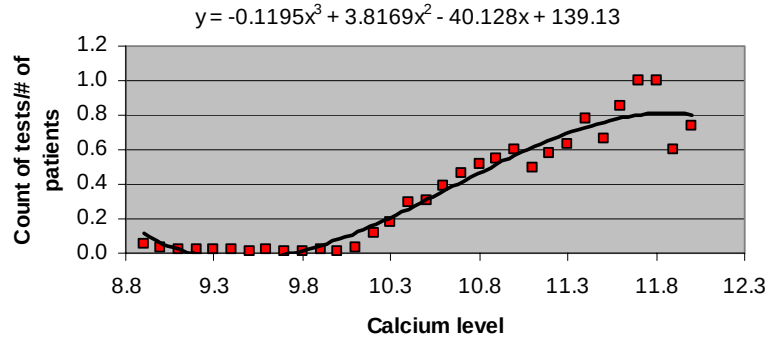


Table 3-3. Assigned Costs per Test or Procedure

Test or Procedure	CPT4 Code	Medicare Per-Unit Reimbursement[a]	Private Payer Per-Unit Reimbursement[b]
Explore Parathyroid Glands	60505	1,378.00	4,031.91
Chest X-Ray	71020	37.11	104.77
Nuclear Scan of Parathyroid	78070	116.70	200.58
Assay Serum Albumin	82040	6.85	18.52
Angiotensin Enzyme Test	82164	20.17	71.22
Assay Calcium in Blood	82310	7.12	18.42
Assay Calcium in Urine	82340	8.34	22.68
Assay Blood Carbon Dioxide	82374	6.76	18.42
Assay Blood Chlorides	82435	6.35	18.42
Assay Cpk in Blood	82550	9.01	18.42
Assay Blood Creatinine	82565	7.07	18.42
Assay Urine Creatinine	82570	7.15	22.68
Assay Ferritin	82728	18.83	51.78
Glucose, Blood,-Gluc. Monitoring Dev.	82962	0.00	13.99
Assay Blood Magnesium	83735	9.26	18.23
Assay of Parathormone (RIA)	83970	57.04	104.16
Assay Alkaline Phosphatase	84075	7.15	16.93
Assay Blood Phosphorus	84100	6.56	16.93
Assay Blood Potassium	84132	6.35	16.93
Assay Blood Sodium	84295	6.65	16.93
Assay Bun	84520	5.45	20.68
Automated Hemogram	85025	10.74	30.56
Prothrombin Time	85610	5.43	23.84
Blood Typing; Abo	86900	4.12	17.81
Culture Specimen, Bacteria	87070	11.90	42.48
Urine Culture, Colony Count	87086	11.16	39.33

[a]Medical reimbursement amounts per test may be lower in practice, if tests are conducted as components of Automated Test Panels (ATPs). About 40 percent of the tests listed in Table 3-3 are reimbursed as ATPs. In the data, of 427,696 of these tests conducted, only 23.1 percent were ordered as a single test (thus not possibly part of a panel).

[b]Weighted averages of 50th percentiles of fees from nine geographic ZIP code regions associated with major medical centers.

lists reimbursement rates for all lab tests by CPT4 code and provides a description of the tests. This file is updated annually. Laboratories may submit higher amounts, but the fee schedule amounts are the full reimbursements for Medicare services.

The private insurance costs were calculated from a weighted average of the 50th percentiles of the fees from nine geographic ZIP code regions associated with major medical centers. This information was obtained from Ingenix.[1]

The reimbursement amounts per procedure listed in Table 3-3 were then used to develop incremental follow-up procedure cost estimates for each patient in the analysis population. As shown Equation (3.1), the individual procedure costs were multiplied by the number of procedures provided in the 12-month period following the initial calcium test:

$$TC_i = \sum (N_{ij} * R_j) \quad (3.1)$$

where

TC_i = total follow-up costs (within 12 months of initial calcium test) for patient i,

N_{ij} = number of occurrences of the jth procedure for patient i, and

R_j = reimbursement rate for the jth procedure.

3.4 IDENTIFYING AND ADJUSTING FOR NONDIAGNOSTIC COSTS

The expected cost functions reflect the correlation between follow-up procedures and calcium values. However, to estimate the change in health care costs associated with measurement bias in calcium test results, it would be ideal to include only those additional procedures that were a direct result of the initial calcium values. These are tests associated with the diagnostic process where hypercalcemia is symptomatic of the disease. However, once an accurate diagnosis has been made (potentially involving extra follow-up procedures due to bias of calcium test results), tests included for the treatment of seriously ill patients should not be included in the economic impact estimates.

For example, intensive care patients may repeatedly receive calcium tests as part of a full blood work-up, and it is possible that some “high-cost” patients may have normal calcium values. Also, patients

[1]Ingenix, Salt Lake City, UT 84116.

diagnosed with hypercalcemia may receive large numbers of calcium tests that are unaffected by potential bias in initial test results. Thus, the final step prior to estimating changes in health care costs due to bias is to adjust the patient cost data to account for health care costs resulting from tests not closely linked to the initial calcium test results.

To account for health care costs not related to the diagnosis of hypercalcemia, an independent data set of 37,817 patients seen at the Mayo Clinic during the first half of 1997 was used.

Subgroups of patients with large numbers of tests were investigated. Some of these patients had large numbers of specific tests ordered during the test period. To limit the influence of these patients on the calcium cost functions (intended to reflect diagnostic testing), limits on the number of follow-up procedures for each test code potentially used in the cost analysis were investigated.

Two percent of the patients account for 19.7 percent of the tests in the 1998–1999 study population, primarily due to intensive monitoring. Because bias in initial calcium tests will predominantly impact follow-up diagnostic tests, patients with large numbers of replicate tests associated with ongoing treatment or monitoring will dilute the cost relationship needed to estimate impacts. This dilution occurs because the monitoring activity is less dependent on the initial test result.

Figure 3-3 shows the distribution of patients by the number of chest X-rays. The curve shows that 64 percent of the population received no follow-up. However, a small subgroup of the population had a large number of tests. For example, 3 percent of the population had four or more chest X-rays following their initial calcium test.

Based on empirical judgment, a limit of three follow-up tests for an individual CPT4 code was used to adjust the costs. Table 3-4 shows the total number of tests and the average number of tests per patient in the study population before and after the adjustment. As shown in Table 3-4, the total number of tests decreased 6.1 percent, from 1.11 billion to 1.04 billion. The average number of tests per patient decreased only slightly, from 12.2 to 11.0, because of the large number of patients with fewer than three tests for individual CPT4 codes. Thus, employing this approach adjusted the number of tests for ongoing-monitoring outliers, without significantly affecting total costs.

Figure 3-3. Cumulative Frequency of Chest X-Rays Ordered per Patient

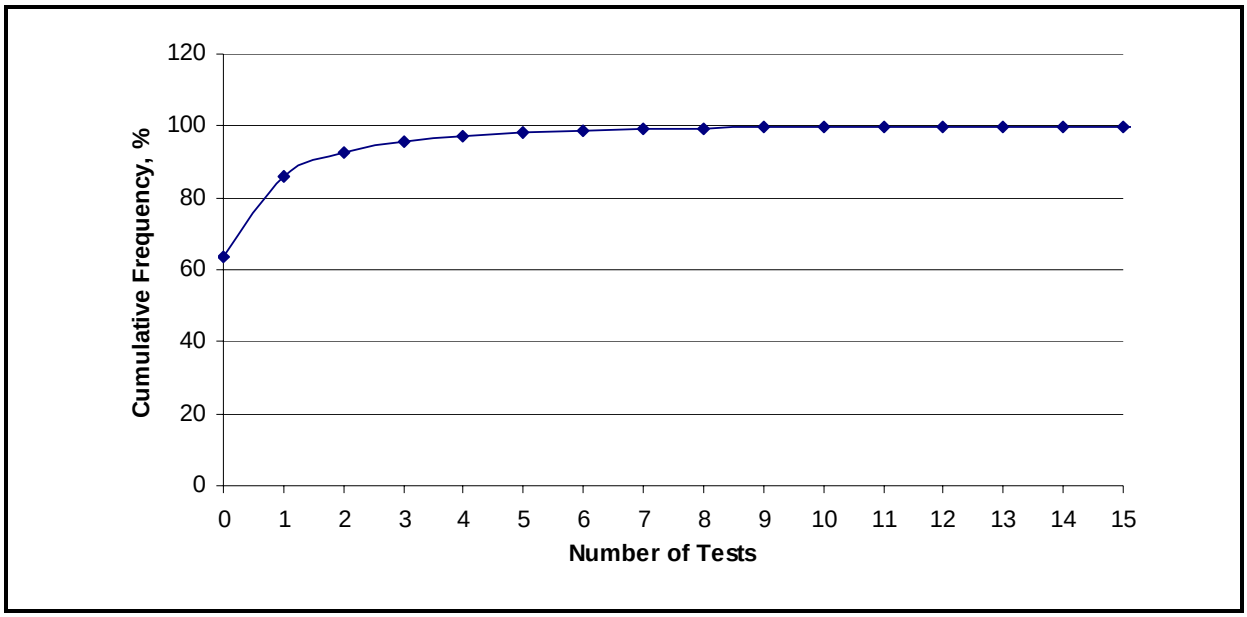


Table 3-4. Number of Tests after Adjustment

	Original Population	Adjusted Population
Total Number of Tests	1,113,182	1,045,107
Average Number of Tests per Patient	12.2	11.0

4 Assessing the Potential Magnitude of Systematic Error

To estimate health care costs associated with systematic calibration error, information is needed on the magnitude of the potential resulting bias in calcium test results. As discussed in Section 2, the size of the bias and the slope of the cost function determine the increase in health care costs.

Based on interviews with industry experts, a range was developed for potential bias in calcium tests resulting from systematic calibration error. Experts identified a range of 0.1 mg/dL to 0.5 mg/dL, and this range is used in Section 5 to estimate potential economic impacts.

4.1 INTERVIEWS

Informal interviews were conducted with four laboratory managers/chief scientists and four equipment manufacturers to investigate the sources of total error (including both random error and systematic error) in calcium test results. The laboratory manager interviews included representatives from independent testing organizations and laboratories affiliated with major hospital systems. The equipment manufacturers produce mostly absorption and emission spectrometry equipment and supply a wide range of clinical and nonclinical testing equipment.

Respondents were initially asked to comment on the sources of uncertainty, focusing on the pre-analytical and analytical phases of testing categories. Within the analytical phase, respondents were asked to distinguish between random error associated with methods, and systematic error associated with traceability, and lot-to-lot variation in calibrators and reagents. Table 4-1 summarizes their responses in terms of the relative importance of each category for calcium test results. For comparison, summary information was requested for cholesterol and prostate specific antigen (PSA) test results, which is also presented in Table 4-1.

Table 4-1. Qualitative Summary of Factors Contributing to Uncertainty

	Calcium	PSA	Cholesterol
Pre-analytical	Very important—handling issues	Minor issue	Very important—diet, activity, etc.
Traceability	Relatively important	Major issue for free PSA; not as important for total PSA	Important, but traceability exists
Methods	Important—different methods can have large offsets	Dependent on antisera characteristics	Important, especially with inexpensive bedside tests
Lot-to-lot variation	Important—manufacturers have trouble with homogeneity	Very problematic, especially at low measurement levels	Related to "bedside devices"; very important

Most respondents indicated that total error was distributed relatively evenly across the calcium testing process and that it was important not to focus solely on a specific aspect, such as methods or traceability. In part, this is because the current cumulative total error of calcium tests is greater than the clinical utility (i.e., recommended upper-bound uncertainty). As one laboratory manager noted, "Of all of the analytes, calcium is the most problematic because of the very tight range for healthy people and the physiological variation is minimal in the population." In general, they acknowledged that systematic calibration error was important, but emphasized that even if this could be completely eliminated, there would still exist significant uncertainty in calcium test results.

As an example, one laboratory manager estimated that the analytical-phase total error for calcium testing was about 5 percent, and that the recommended upper bound is about 3.3 percent. The implication is that the marginal benefits from reducing the error of individual testing components or phases may be small unless total error can be reduced relative to the "normal" human calcium reference range.

4.2 PRE-ANALYTICAL PHASE

Respondents generally believed that the pre-analytical phase is responsible for about one-quarter to one-half of the total error associated with laboratory test results for hypercalcemia. The handling of blood samples is the main source of uncertainty. Issues cited were

- human error,
- change in pH that can result from air bubbles entering the sample,
- choice of anticoagulate and how it binds with calcium, and
- clotting because of freezing and thawing.

For example, clotting problems occasionally exist with dialysis patients requesting PTH tests because of samples sitting and/or the freezing and thawing process. Filters can be used to remove fiber clots, but sometimes test results are still unreliable.

4.3 ANALYTICAL PHASE

Respondents generally agreed that the analytical phase accounts for at least half of the uncertainty introduced into calcium test results. Three main factors (ranked in order of importance) mentioned were

- methods used by the laboratory instruments,
- lot-to-lot variations in reagents and calibrators, and
- traceability of reference material.

Different methods used in analyzers and lot-to-lot variations in calibrators and reagents were cited as approximately equally important sources of systematic error introduced in the analytical phase. In general, the offsets between different analyzers that use different methods can be large and can lead to biased results from the equipment. A commonly cited shortcoming is that testing is rarely conducted for bias/accuracy. Most equipment is only tested for precision to meet FDA requirements. FDA does not require certification or reporting of equipment accuracy.

Respondents estimated that offsets range up to 0.5 mg/dL and result from differences between methods and from poor manufacturing quality control for analyzers using the same methods. In addition, controls read differently on different analyzers, and controls used in laboratories are not used for accuracy tests. As a result, test results may not transfer from one integrated delivery system (IDS) or physician to another, resulting in physicians' commonly establishing each patient's own baseline with test results from a failure source.

Respondents said that laboratories and physicians often compensate systematic error in equipment offsets by establishing their own baselines and repeatedly using the same laboratories and/or the same suppliers of instruments and reagents. Although this approach ensures consistency over time, it can generate excess testing costs and is only practical for

large IDSs with nontransient patient populations. A related outcome is reduced competition and customer lock-in to particular laboratories or equipment/reagent manufacturers.

Traceability was mentioned by all respondents as an important source of systematic error in the analytical phase. However, they disagreed somewhat on the relative importance of traceability, given the other factors contributing to total error in calcium test results.

Human error was also mentioned as a potential issue in the analytical phase when new or inexperienced employees are involved in mixing reagents or calibrators. However, human errors are likely to be isolated events, leading to easily identifiable “bad” test results (and retesting), and do not lead to systematic errors in test results.

Finally, matrix effects were also mentioned as factors contributing to uncertainty. For example, a device may be well correlated to blood but not to the calibrator solution.

5 Economic Impacts of Systematic Error in Calcium Measurements

Systematic error leading to analytic measurement bias for serum calcium will shift the cost functions developed in Section 3. This section illustrates the economic impacts for different levels of analytical bias ranging from 0.1 to 0.5 mg/dL. It is estimated that the increased health care costs associated with these shifts range from $8 to $31 per patient, which translates into a national increase in health care costs of approximately $60 to $199 million per year.

5.1 CHANGE IN HEALTH CARE COSTS PER PATIENT

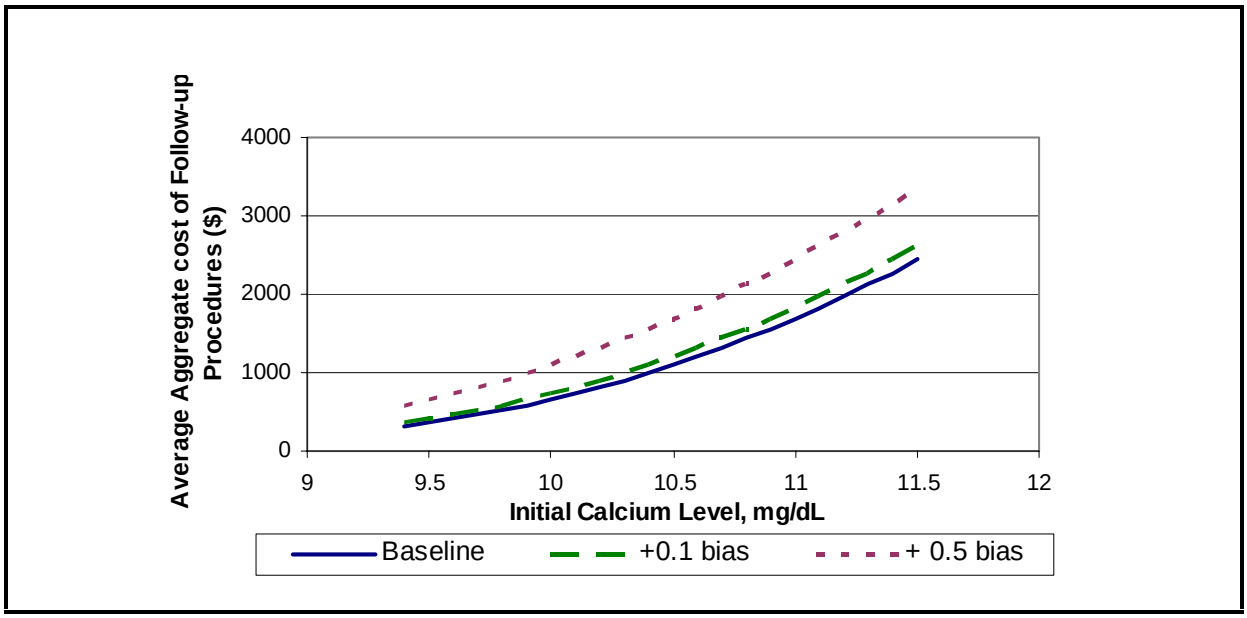
As shown in Figure 5-1, health care costs resulting from follow-up procedures are an increasing function of the initial calcium test value. The average per patient follow-up costs are about $650 for patients with an initial calcium value of 10.0 and $1,700 for patients with an initial calcium value of 11.0.

The impact of analytical bias can be simulated as an upward shift in the per-patient cost curve. For example, if a 0.5 mg/dL bias results in a patient receiving a test result of 11.5 mg/dL instead of 11.0 mg/dL, this would on average lead to $700 in additional follow-up tests. The shift is smaller at lower calcium values.[1] A patient receiving a test result of 10.5 mg/dL instead of 10.0 would on average receive an additional $550 in follow-up tests. This upward shift in the cost function is also shown in Figure 5-1. The incremental cost associated with bias is a function of the slope of the cost curve in Figure 5-1.[2] Costs are expressed per patient and capture all relevant follow-up procedures.

[1] A particular bias is assumed to be the same across all calcium levels.

[2] Note that because most initial calcium values are less than 10.5 (as shown in Figure 3-1) and normal ranges have lower incremental costs, the average incremental cost per patient becomes relatively small.

Figure 5-1. Shift in the Cost Function due to Analytic Bias
Private insurance patients



An alterative presentation of this impact is illustrated in Figure 5-2. Separate curves are shown representing the average change in health care costs per patient for different initial calcium levels as analytic bias increases. For calcium values of 9.9 to 10.0 mg/dL, minimal impact is associated with bias because, even with the potential bias, the test results are still relatively close to the reference range. However, as the initial calcium value becomes elevated, the impact of analytic bias becomes large as health care costs due to follow-up diagnostic tests significantly increase.

As shown in Figure 5-2, the patterns are similar for all four gender-payment subgroups. However, female patients had larger changes than male patients, and the percent changes were larger for private insurance compared to Medicare patients for both genders. The figures illustrate that both positive and negative shifts increase costs. However, because the model was built to represent the effects of hypercalcemia, the positive shifts are likely to be more accurately represented than negative shifts.

Figure 5-2. Cost Impact of Analytic Bias for Different Ranges of Calcium Values

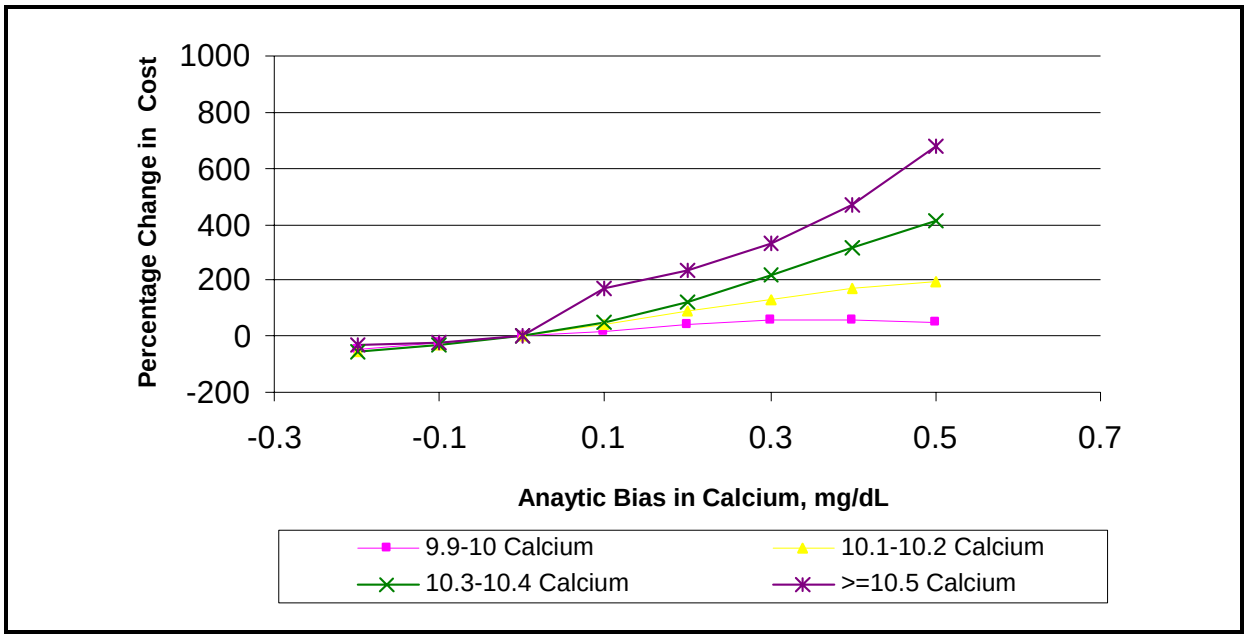


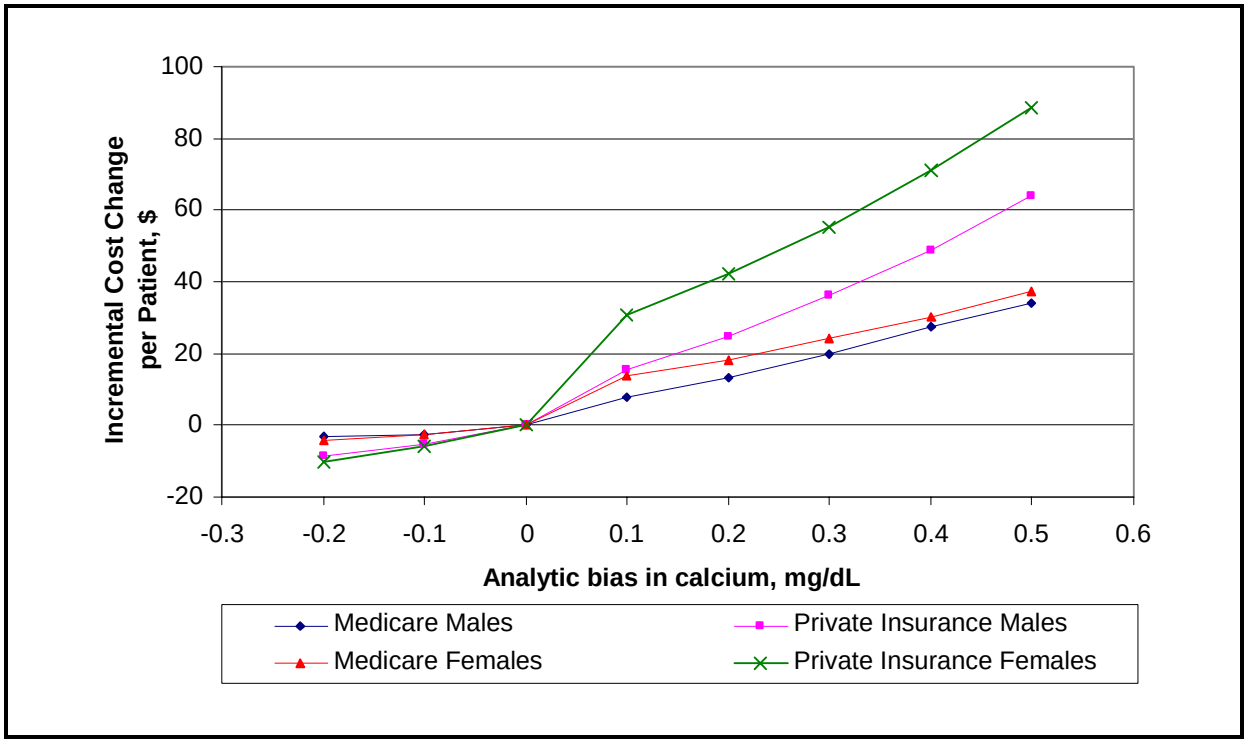
Table 5-1 presents incremental cost estimates for patients with calcium values greater than or equal to 8.9 mg/dL. Patients with calcium values of less than 8.9 mg/dL were excluded from this table (and the process used to estimate national impacts) because these patients were below the acceptable reference range and the cost functions were developed to investigate cost associated with hypercalcemia.

Table 5-1. Incremental Costs per Patient (≥ 8.9 mg/dL)

	Analytic Bias: Cost per Patient	
Sub Segments	**0.1 mg/dL**	**0.5 mg/dL**
Private Insurance		
Male	$15.2	$63.8
Female	$30.8	$88.6
Weighted Average	$23.0	$76.2
Medicare		
Male	$7.8	$34.2
Female	$13.7	$37.5
Weighted Average	$10.8	$35.9
Total Weighted Average	$16.9	$56.0

As shown in Figure 5-3, private insurance patients' cost increases are estimated to be 2 to 3 times as great as Medicare patients' cost increases. This estimate is primarily driven by the differences in the reimbursement rates provided in Table 3-1.

Figure 5-3. Average Cost Impact of Analytic Bias for Patients with Calcium Values ≥ 8.9 mg/dL



The range between private insurance and Medicare reimbursement rates provides an approximate upper and lower bound for incremental health care costs associated with bias.[3] As shown in Table 5-1, average incremental costs for analytic bias of 0.1 mg/dL range from $8 to $31 based on Medicare and private insurance reimbursement rates. For analytic bias of 0.5 mg/dL, incremental costs range from $34 to $89 per patient.

5.2 IMPACT ESTIMATES

Table 5-1 presents incremental costs per patient. To estimate the national impact of systematic calibration error, the number of patients affected each year needs to be determined. It is not realistic to assume

[3]It should also be noted that reimbursement rates do not exactly reflect the increased costs to society. In many instances, Medicare costs do not fully compensate hospitals or laboratories for their operating and material costs. In addition, private insurance rates may include profits that are transfer payments and not social costs.

that all, or even a large share of patients are affected because if measurement error were widespread and persistent, medical decision making would adjust, incorporating the bias into the baseline.

For this analysis, we simulate a scenario where a major equipment manufacturer distributes a "lot" of calibrators with an undetected systematic error in the reference material. This results in analytic bias in calcium tests at all laboratories supplied by this manufacturer. It is not likely that multiple manufacturers would issue calibrator lots with bias at the same time because these are sporadic events. Thus, the market share for a representative equipment manufacturer is used as the affected population when quantifying the potential impact of bias.

Table 5-2 presents the market share of major equipment manufacturers of laboratory test equipment. The largest three suppliers are Dade Behring, Inc., Beckman Coulter, Inc., and Ortho-Clinical Diagnostics, with each holding approximately one-fourth of the market. Because equipment manufacturers also produce and distribute calibrator and reagents for their equipment, it is plausible that a single manufacturer could impact 15 percent of calcium test for up to 1 year. This assumption of 15 percent is used in estimating the potentially affected population for the impact analysis.

Table 5-2. Market Share for Chemistry Instrument Installed Base: Hospital Labs

Instrument Manufacturers	1997[a]	2001
Dade Behring	26%	28%
Beckman Coulter	22%	25%
Ortho-Clinical	18%	21%
Roche Diagnostics	18%	14%
Olympus	—[b]	4%
Bayer Diagnostics	3%	3%
Abbott Diagnostics	5%	2%
Other	12%	3%
Totals	**100%**	**100%**

Source: IMV Ltd., 2002. LABSTAT Instrument Report: Automated Chemistry Analyzers Year-End 2001. IMV Ltd.: Greenbelt, MD.

[a]The market shares for installed instrument base in 1997 included commercial labs. However, the 2001 installed base does not include commercial labs.

[b]In 1997, Olympus was included in the "Other" category.

5.2.1 Population Weights

Incremental costs are based on patients with the following characteristics:

- age 18 and older and
- having an initial calcium test result greater than 8.9 mg/dL.

To develop an appropriate population weight associated with national impacts, the number of patients in 1998 meeting the above characteristics from Olmsted County in Minnesota was compared to the adult population of Olmsted County. This ratio, based on geography, is used to determine the share of the population for which incremental health care costs are appropriate. This share (i.e., the proportion of the population receiving an initial calcium test each year) is then applied to the U.S. population to estimate the "affected population."

Other weights were investigated, such as the ratio of the number of initial calcium tests to the total number of patients admitted to Mayo Clinic. However, it was determined that because Mayo Clinic is a referral hospital, the number of tests received by its patients may not be representative of the national health care system as a whole. For this reason, a geographic weighting scheme was preferred as opposed to one based on Mayo Clinic's patient population.

It is estimated that approximately 23.7 million patients, aged 18 and older, have an initial calcium test result greater than 8.9 mg/dL each year (based on 1998 Mayo Clinic and U.S. census data), and an error introduced by a single large instrument manufacturer could potentially affect 15 percent of the tests. Table 5-3 summarizes the calculation.

5.2.2 National Impact Estimates

The incremental costs per patient presented in Table 5-1 and the population weight were used to develop a rough estimate of the potential impact of analytic bias on U.S. health care costs in 2000 (see Table 5-4). Using private insurance and medical reimbursement rates for the year 2000 as the upper and lower bounds, respectively, economic impacts are estimated to range from $38.3 to $81.6 million for an analytic bias of 0.1 mg/dL and from $127.4 to $270.5 million for an analytic bias of 0.5 mg/dL.

Table 5-3. Population Data Used in Cost Extrapolation

Adult patients at Mayo Clinic in 1998 from Olmsted county with initial calcium test values greater than 8.9 mg/dL	9,611
Adult population in Olmsted County in 1998[a]	83,700
Share of adult population	11.4%
U.S. adult population in 2000[b]	206.7 million
Population receiving calcium tests in a year (11.4 percent of U.S. adult population)	23.7 million
Share of tests affected by systematic error[c]	15%
Number of Patients Affected by Bias	**3.5 million**

[a]Source: U.S. Census population estimates, ESRI.

[b]Source: U.S. Census, 2000.

[c] Represents potential error introduced by a large instrument manufacturer.

Table 5-4. National Cost Estimates (based on 3.5 million patients)

	Analytic Bias: National Cost ($Millions)	
Sub Segments	**0.1 mg/dL**	**0.5 mg/dL**
Private Insurance	82	271
Medicare	38	127
Total Weighted Average	60	199

Using the total weighted average cost shown in the last row of Table 5-1, the economic impact estimates for potential bias range from $60 to $199 for analytic bias of 0.1 to 0.5 mg/dL, respectively.

5.3 REPEATED TESTS

Uncertainty in laboratory test results can affect not only the frequency of procedures ordered by physicians but also the number of "initial" calcium tests themselves. These cost inefficiencies are not included in the quantitative impact estimates presented above, but are discussed qualitatively in this section.

5.3.1 Replicated Tests

As patients increasingly move between IDSs as a result of increased competition between health care providers, the transferability of test results becomes an important issue. It has been suggested that many

laboratory tests ordered are repeated because physicians are unfamiliar with different methods or believe that they are not comparable.

Hospitals have implemented electronic systems to link patient records in an effort to reduce unnecessary duplication of lab tests. But these systems have reduced the ordering of "redundant" lab tests by only 9 to 15 percent (Bates et al., 1998; Van Loon et al., 1999). An interesting question is why the remaining 85 percent of seemingly repeat tests persist. Despite recent establishment concerns about excessive testing, test ordering propensity has been surprisingly resistant to change.

One study by Bates et al. (1999) closely examined physician test ordering behavior in a randomized control trial setting, conducting careful chart review to scrutinize whether apparently repeated tests were actually justified on medical grounds. Only 41 percent of the repeated tests appeared to be justified.

This study suggests that apparent redundancy may be caused by uncertainty due to poor test quality. In this scenario, ordering duplicate tests can be seen as a strategy to reduce this uncertainty—especially if some time has elapsed since the first test or if it is likely that another lab will conduct the test, which would yield a second, independent observation on the test result. The evidence cited above suggests that this "quality bias" toward repeat testing may be quite large. If as many as half of all repeat tests are due to uncertainty, then the direct cost impact on society is huge. One large New York hospital saved $632,000 in a single year by installing an electronic checking system for duplicate tests (Nemes, 2002). In addition to these direct costs, repeat test ordering can lead to more false-positive results, which can lead to further unnecessary treatment and increased costs (Bates, Goldman, and Lee, 1991). Thus, the cost impact on society of poor quality standards in laboratory testing can be very large.

5.3.2 Evidence of Replicated Tests

For this study, a small data set of laboratory test claims was obtained from Tom Johnson at Blue Cross/Blue Shield of Kansas, in Topeka, Kansas. A total of 474 patients were included, of which 202 patients had only one physician provider and 272 patients had multiple physician providers in 1999 and 2000. Twelve percent of the patients seeing only one provider had at least one replicate test requested, whereas 51 percent of the patients seeing multiple providers had at least one replicate test. Part of this difference in repeat testing is related to the

higher number of physician encounters for the multiple physician group (mean 38, median 7), compared to the single physician provider group (mean 2.4, median 1). The provider laboratories were not identified in this data set; however, the markedly higher number of patients with replicate tests in the multiple physician data set suggests that some of the additional testing may have been related to uncertainty in test results, potentially related to differences between laboratories.

References

Atomic Spectroscopy. 2002. <http://elchem.kaist.ac.kr/vt/chem-ed/spec/atomic>. As obtained October 28, 2002.

Bates, D., L. Goldman, and T. Lee. 1991. "Contaminant Blood Cultures and Resource Utilization. The True Consequences of False-Positive Results." *Journal of the American Medical Association* 265:365-369.

Bates, D., G. Kuperman, D. Boyle, et al. 1998. "Clinical Laboratory Tests: What Proportion are Redundant?" *American Journal of Medicine* 104:361-368.

Bates, D., G. Kuperman, E. Rittenberg, et al. 1999. "A Randomized Trial of a Computer-based Intervention to Reduce Utilization of Redundant Laboratory Tests." *American Journal of Medicine* 106:144-150.

Chem USA. <www.chem.usu.edu/faculty/sbialkow/classes/3610/flame.html>. As obtained October 28, 2002.

Harvard Medical School (HMS). July 2001. Consumer Health Information *InteliHealth*. <http://www.intelihealth.com>; search words "parathyroid cancer."

Klee, G., P. Kao, and H. Heath. 1988. "Hypercalcemia." *Endocrinology and Metabolism Clinics of North America* 17(3):September.

Nemes, J. 2002. "Hospitals Seek a Tech Cure for Mistakes, Red Tape." *Crain's New York Business*, <http:/www.crainsny.com/page.cms?pageID=408>. As obtained October 28, 2002.

Petersen, P., C. Verdier, T. Groth, C. Fraser, O. Blaabjerg, and M. Horder. "The Influence of Analytical Bias on Diagnostic Misclassifications." *Clinica Chimica Acta* 260: 189-206.

U.S. Census. 2000. <http://www.census.gov/Press-Release/www/2002/sumfile3.html>. As obtained on March 12, 2003.

Van Loon, Herman, Frank Bunting, Jan Herman, and Robert Van den Over. 1999. "The Frequency of Meaningless Repeated Lab Tests in Belgian Medicine." First European Network Organisations Open Conference, WONCA 1999. <http:/www.equip.ch/e_1/meetings/docmal/doc_7.html>. As obtained October 28, 2002.

A Reference Methods

The absorbance ratio reflects the relative absorption rates of the sample versus the calibrator. Because the absorption factor is used to directly adjust the calcium reading, any errors in measuring the absorbance can affect test results. Therefore, any imprecision (random error) in the absorbance measurements contributes to the uncertainty (total error) regarding the calcium concentration.

Atomic spectroscopy (also referred to as atomic spectrometry) is the method of choice for measuring calcium (and most other elements) because of its accuracy (Atomic Spectroscopy, 2002; Klee, Kao, and Heath, 1988). However, most automated laboratory test equipment uses photometry because this method can be used to acceptably screen large numbers of specimens at lower costs. Table A-1 provides a description of the primary blood and urine tests available for detecting hypercalcemia and hyperparathyroidism. Appendix B describes the common laboratory test equipment along with an overview of equipment manufactures.

A.1 Photometry

Photometry is an analytical technique used to measure the color produced by the interaction of calcium with a dyelike substance such as cresolphthalien. A spectrophotometer is used to measure the amount of light that a sample absorbs. The instrument operates by passing a beam of light through a sample and measuring the intensity of light reaching a detector.

Table A-1. Various Analytic Procedures for Determining Serum Calcium Levels

Method	Type of Calcium Detected	Description
Photometry	Total calcium	At a pH of 10 to 12, calcium yields a red complex with orthocresolphthalien-complexone. Other additives are used to eliminate interference. Photometry is used in automated multichannel clinical chemistry analyzers.
Atomic Absorption Spectrometry	Total calcium	Calcium in serum or urine is diluted with lanthanum chloride solution to bond interfering substances such as proteins and phosphates. When the solution is introduced to a flame, certain wavelengths emerge indicating the presence of calcium. This method's imprecision between series ranges from 1 to 2 percent.
Atomic Emission Spectroscopy	Total calcium	Serum is diluted with distilled water, sprayed into a flame of acetylene, and vaporized. Simultaneously, emissions of calcium may be measured.
Fluorometric Titration	Total calcium	This method is used in several analyzers. However, it is susceptible to interference by copper, iron, zinc, and certain drugs.

The light, emitting a constant number of photons per second, passes through the analyte (the substance being measured), and some of the photons are absorbed. The absorption of photons reduces the intensity of the light, and the effect is measured by a detector on the opposite side. The absorption rate is then used to calculate the concentration level of the analyte.

Calibration for photometry tests typically consists of "zeroing" the photometer using distilled water as a blank and establishing an upper end measurement with a high concentration solution of the chemical of interest (e.g., 100 µg/mL). The emissions intensity of three or four other standard concentration solutions are then measured at incrementally lower levels. This process is repeated several times to check the accuracy of the system (Chem USA, 2002).

A.2 Atomic Spectroscopy

Atomic spectroscopy is commonly segmented into atomic absorption spectroscopy (AAS) and atomic emission spectroscopy (AES). The underlying principle of all atomic methods of analysis is that the sample be decomposed to the greatest extent possible into constituent atoms. The gas-phase atomic cloud is then analyzed using ultraviolet, visible, and near-infrared regions of the electromagnetic spectrum.

Atomic absorption spectroscopy uses the absorption of light to measure the concentration of atoms in the gas. The disadvantage of AAS is that it is difficult to measure more than one element at a time because of the generation of a broad-band spectral background resulting from residual molecules in the atomic source or from smoke generated from the atomic formation process.

In contrast, AES measures the optical emissions from excited atoms to determine concentrations. As part of the atomization process, high-temperature gases are created with sufficient energy to provoke the atoms into high energy levels. When the atoms decay back to lower energy levels, they emit their signature spectrums. AES is a multi-element procedure, making it possible to perform simultaneous multi-element determinations using multichannel detection systems.

In all atomic spectroscopy, the degree of atomization is an important component of instrument sensitivity. Less than complete atomization results in lower sensitivities in the atomic method. However, even more detrimental are variations in the fraction of atomization (matrix effects), because variations in the extent of atomization from sample to sample or from sample to standard will lead to errors in calibration.

B Manufacturers of Test Equipment

The supply chain for calcium testing equipment includes two main groups of companies. The first group includes the suppliers of the technology, equipment, and reagents. The second group is the producers of reference materials. In many instances, the same company provides both.

B.1 MANUFACTURERS OF TESTING EQUIPMENT AND REFERENCE MATERIALS

Numerous types of companies produce technologies that can be used in the calcium testing process. Table B-1 lists (by equipment type) companies that produce calcium testing equipment.

The companies that produce the reference materials used to calibrate tests are as important as the companies that produce testing equipment and reagents. To sell standard reference materials for calcium, a company has to achieve a 510(k) ranking from the U.S. Food and Drug Administration (FDA). This ranking allows the traceability of the reference material. Table B-2 lists companies that produce reference materials for calcium testing.

Table B-1. Producers of Calcium Testing Equipment

Spectrophotometric Analysis/Colorimetric Test Equipment Producers	
Abbott Laboratories	EMD Chemicals (formerly EM Science)
Bausch & Lomb	Hach Company
Bayer Diagnostics	Johnson & Johnson
Beckman Coulter Inc.	Macherey-Nagel Inc.
BMD Hitachi	Ocean Optics OEM
Buck Scientific	Ortho Clinical Diagnostics
Ciba-Corning	PerkinElmer Instruments
Dade Behring Inc.	Roche Diagnostics
Dupont	Thermo Orion Corporation
Atomic Absorption Spectrometry Equipment Producers	
Analytik Jena AG	Multichannel Instruments AB
Anglia Instruments Ltd	PerkinElmer Instruments
Aurora Instruments Ltd.	Solent Scientific Ltd
Cathodean Ltd	Spectrolab Analytical
CETAC Technologies	Thermo Elemental
GBC Scientific Equipment	Varsal Instruments, Inc.
Infometrix Inc.	
Flame Atomic Emission Spectroscopy Equipment Producers	
Aurora Instruments Ltd.	Leco Corporation
Agilent Technologies Inc.	Leeman Labs Germany GmbH
Anglia Instruments Ltd	Multichannel Instruments AB
Automated Fusion Technology	Spectro Analytical Instruments
GBC Scientific Equipment	Thermo Orion Corporation
JY Horiba	Thermo Elemental
Ion Selective Electrodes (ISE for Ionized Calcium)	
Beckman Coulter (UK) Ltd	Qcl Ltd
Mettler-Toledo Ltd	Vernier Software & Technology
Thermo Orion Corporation	

Table B-2. Producers of Reference Materials

Reagent/Standardized Reference Material Producers	
LGC	Teco Diagnostics
LaMotte Company	OFI Testing Equipment, Inc.
BIOTREND Chemikalien GmbH	

B.2 LEADING MANUFACTURERS OF HIGH-VOLUME LABORATORY TESTING EQUIPMENT

Abbott Laboratories[1]
(Source: http://www.abbottdiagnostics.com/our_division/index.htm)

Originally founded as a pharmaceutical medicine laboratory in 1900, Abbott Labs has become one of the largest diagnostic equipment designers in the world. Abbott employs over 60,000 people, 5,000 of which are research scientists involved in developing new technologies and products. The company's research fields are in the areas of diagnostics and immunodiagnostics, hematology, blood glucose monitoring, and DNA testing.

The diagnostic division of Abbott Labs recently acquired Vysis Inc., a leading genomic disease management company specializing in clinical laboratory equipment. Their most recent chemistry analyzer system, the Architect *i2000* came on the market in 1999.

The Architect *i2000* conducts immunoassays using chemiluminescence detection technology and can perform hundreds of tests per hour with up to 25 reagents within the system.

Bechman Coulter, Inc.
(Source: http://www.beckman.com/products/instrument/genchem/lx2000pro.asp)

Beckman Coulter makes products used in hospital laboratories, physicians' offices, and group practices. The company provides a variety of systems for medical research, drug discovery, and biotechnology applications. The company recorded $2 billion in sales for FY2001 and employs nearly 10,000 people. Bechman acquired Coulter Corp. in 1997, which allowed the company to offer a comprehensive

[1]Average cost of systems listed range from approximately $100,000 to $150,000 and all perform calcium testing (see test menus for further information).

product listing that spans the fields of life sciences, clinical diagnostics, and cellular analysis.

Synchron System is the clinical laboratory product line, offering automated clinical chemistry tests that include calcium testing. Pictured below is the Synchron LX *20PRO,* which offers closed-tube sampling and a Near Infrared Particle Immunoassay (NIPIA) detection system and is capable of conducting 1,540 tests per hour.

Dade Behring, Inc.
(Source: http://www.dadebehring.com/edbna2/ebusiness/home.jsp?lang=E)

Dade International and Behring Diagnostics merged in 1997 to form one of the largest diagnostic companies in the world, employing over 6,500 people worldwide and generating 1.2 billion in revenue for FY2001. Their business is dedicated entirely to diagnostics.

The Dimension, Dade Behring's product line of integrated chemistry systems, provides automated testing techniques for small labs with low- to medium-volume testing. The system offers reagent management to support calibration procedures for routine chemistries. The company offers this type of testing equipment in a variety of sizes tailored specifically to the need for different volumes of testing.

Ortho-Clinical Diagnostics
(Source: http://www.jnjgateway.com/home.jhtml?loc=USENG&spec=allSpecialties)

Ortho-Clinical Diagnostics, a Johnson & Johnson company, provides diagnostic products and services for the health care sector.

The VITROS 950, pictured below, provides high-volume testing capabilities. With throughput of up to 900 results per hour, the VITROS 950 can store and perform a full complement of VITROS chemistry assays on a continuous basis. The VITROS 950 can also be automated and/or incorporated into a workcell or laboratory automation system.

Roche Diagnostic
(Source: http://www.roche-diagnostics.com)

Roche Diagnostic is a division of F. Hoffmann-La Roche Ltd, based in Basel, Switzerland. The company employs more that 16,500 people and reported $2.6 billion in sales for FY2001. In 1991, Roche released the first fully automated immunochemistry system, named the "Cobas Core," and in 2002 the company unveiled its most advanced version of the Cobas line with the "Cobas Integra 400*plus*."

The Cobas Integra 400*plus,* shown below, is capable of 400 tests per hour, using four on-board measurement technologies and robotic handling of reaction cuvettes. The system supports a broad menu of tests including calcium testing.