REPORT BY THE
AUDITOR GENERAL
OF CALIFORNIA

COST EFFECTIVENESS OF THE MEDI-CAL THERAPEUTIC DRUG UTILIZATION REVIEW PROGRAM

C-775 MAY 1991

Kurt R. Sjoberg, Auditor General (acting)

State of California
Office of the Auditor General
660 J Street, Suite 300, Sacramento, CA 95814
Telephone : (916) 445-0255

May 3, 1991 C-775

Honorable Robert J. Campbell, Chairman
Members, Joint Legislative Audit Committee
State Capitol, Room 2163
Sacramento, California 95814

The Office of the Auditor General presents a report prepared under contract by Ernst & Young concerning the cost effectiveness of the pilot therapeutic drug utilization review program.

This audit was conducted to comply with Chapter 1340, Statutes of 1987.

Respectfully submitted,

KURT R. SJOBERG
Auditor General (acting)

STATE OF CALIFORNIA

OFFICE OF THE AUDITOR GENERAL

Cost-Effectiveness of the Medi-Cal Therapeutic Drug Utilization Review Program

ERNST & YOUNG

COST-EFFECTIVENESS OF THE MEDI-CAL THERAPEUTIC DRUG UTILIZATION REVIEW (TDUR) PROGRAM

TABLE OF CONTENTS

APPENDICES

SUMMARY

ERNST & YOUNG

SUMMARY

Results In Brief

The California Office of the Auditor General, in compliance with state legislation, contracted with Ernst & Young to evaluate the state's pilot Medi-Cal Therapeutic Drug Utilization Review (TDUR) program which is operated by the California Department of Health Services (DHS). The objective of the evaluation was to determine the cost-effectiveness of TDUR and to develop recommendations regarding continuation of the program. The purpose of this report is to document our findings and conclusions with respect to these objectives. We determined that the pilot TDUR program had a direct impact resulting in decreased utilization of drugs, outpatient services and hospital care for a small group of Medi-Cal recipients during the review period from July 1, 1989 through August 30, 1990. However, the direct cost savings associated with the reduced service utilization were too small to prove the cost-effectiveness of the program. We found that costs exceeded benefits by approximately $1.5 million. We did not measure presumed indirect benefits of the pilot; however, to be cost-effective, indirect benefits would have to be in the range of 13 to 24 times the direct benefits. Based on our findings, we do not believe the pilot program has been cost-effective, but we believe opportunities exist for substantially increasing measurable program benefits.

Background

Many in the medical community believe that a TDUR program is an effective means of improving the health of Medicaid eligibles so that the incidence of institutionalization is reduced. As a result, health care costs are expected to decline under such programs. The intent of California's pilot TDUR program was to test this hypothesis. An experiment was designed so that health care costs for Medi-Cal eligibles could be compared for a test group receiving TDUR intervention in three counties, versus a control group in two counties for whom the intervention did not occur.

The TDUR process involved monthly computer analysis of recipient Medi-Cal claims for drugs, outpatient services, and inpatient care. The computer selected recipients with atypical claims histories for further review. Committees of medical professionals then selected a portion of the recipients for follow-up notification and subsequent re-review. Those selected by the committees for follow-up activities represented the eligibles receiving the TDUR intervention, and for whom direct program benefits should have accrued.

Summary of Findings

- *The California TDUR Program Reduced Utilization of Pharmacy,Outpatient, and Inpatient Services for Medi-Cal Eligibles in Pilot Counties*

  Our analysis confirmed that a portion of the patients selected for TDUR intervention showed improvement in therapeutic care through a combination of reduced drug prescriptions and more appropriate drug regimens. Individuals whose drug therapies improved during the pilot experienced an associated decrease in health care costs. However, improvements were not noted for long-term care patients. We do not attribute this outcome to the TDUR process. Rather, the 90-day medical profiles used in our analysis did not afford sufficient data to determine changes for these individuals, who are typically institutionalized for an average of 280 days.

- *Based on Measured Direct Savings, The Pilot TDUR Program Was Not Cost Beneficial*

  Pilot program costs exceeded maximum direct benefits by $1,598,409. We did not measure presumed indirect benefits. However, our findings indicated that the pilot was not cost-effective, and that the TDUR program, *as it is presently operated,* would not be cost beneficial for the State of California.

- *Implementing a Cost-Effective Statewide TDUR Program in California Will Require Significant But Potentially Feasible Changes from the Pilot Program*

  We believe the potential exists for a TDUR program to provide an effective means of controlling Medi-Cal costs in California. The scope of this study did not include evaluating alternative types of programs, such as prospective TDUR. These are options the state may wish to consider. However, if the current retrospective TDUR process is retained, we found a number of ways to make the program more cost-effective.

Recommendations

To improve the cost-effectiveness of its current retrospective TDUR program, the state should ensure that the following changes are implemented:

- New exceptions criteria should be developed for the TDUR process. This improvement possibly offers the greatest potential for increased savings
- Exceptions criteria should be applied more selectively
- The Department of Health Services should provide more rigorous administration of the TDUR contract
- Master TDUR Committee members should develop relationships with county review committees and elicit feedback from them regarding program operations
- Local review committees should be given the ability to follow up on cases selected for intervention, if needed
- Lab test data should be included as part of the information available to local review committees
- The accuracy and completeness of the state's provider files should be improved.

Agency Comments

The Department of Health Services does not concur with the findings of this study with respect to cost-effectiveness. The Department does not offer any conclusions of its own. Instead, it recommends using a report which is scheduled for publication on May 24, 1991. This report will be issued by Stanford Research Institute, a subcontractor to the Department's contractor, Virginia Computer Company.

Although the Department does not concur with the study findings, it acknowledges the value of the recommendations to improve the cost-effectiveness of the TDUR program. The Department says these recommendations will be considered in planning the statewide Drug Utilization Review program mandated by the federal Omnibus Budget Reconciliation Act of 1990.

INTRODUCTION

ERNST & YOUNG

INTRODUCTION

Summary of Legislative Mandate

The State's legislative mandate for the TDUR pilot project derives primarily from Assembly Bill No. 2606, Chapter 1340, Statutes of 1987. This legislation established a pilot open formulary and drug utilization review project in the Medi-Cal program. A formulary is simply a list of drugs. California has traditionally utilized a pseudo Closed Formulary approach to monitor drug usage and contain costs within the Medi-Cal program. A Closed Formulary system means that no drugs may be prescribed that are not on the list. The California Medi-Cal program has a *pseudo* Closed Formulary system because drugs may not be prescribed if they do not appear on the list *unless prior authorization is obtained.* On the other hand, an Open Formulary system allows all drugs to be prescribed, and no prior authorization by Medi-Cal would be required.

Legislation leading up to AB 2606 included AB 388 and AB 2655, approved in August, 1984. AB 388 was a statute to institute a statewide DUR program. AB 2655 provided for an experimental Open Formulary drug project to be conducted in Sacramento and Placer counties. These two bills were subsequently combined for administrative reasons into AB 73 in December, 1984. However, implementation of the DUR program and Open Formulary experiment were dependent on obtaining a federal waiver to allow the Open Formulary in just two counties. In 1984, federal law required that all Medi-Cal recipients in a state be subject to a single policy regarding formularies. Therefore, the federal waiver was not granted to test the Open Formulary concept in just two counties, and neither the DUR Program nor the Open Formulary experiment could be implemented as intended in the legislation.

AB 2606 was written to allow the implementation of TDUR independent of the Open Formulary experiment. A federal waiver was required to implement the open formulary component of the legislation. This waiver was requested and was subsequently denied by the federal government. As a result, the open formulary experiment could not be implemented. The implementation of the remainder of the legislation, including TDUR, proceeded. TDUR is currently operated in a closed formulary/prior authorization environment. This study covers the portions of AB 2606 relevant to the TDUR implementation.

Chapter 456, Statutes of 1990, allowed California to change from a pseudo Closed Formulary to a list of contract drugs. Only drugs whose manufacturers have a contract with the state are included on the list. The implementation of this change was accomplished on a statewide basis. This means that it was implemented concurrently in all California counties including those counties a part of the TDUR pilot program.

As described in AB 2606, the law established a "...therapeutic drug utilization review system for the evaluation of therapeutic outcomes of drugs prescribed for persons eligible for Medi-Cal..." This bill authorized the Department of Health Services (DHS) to competitively procure the services of an expert contractor to implement a therapeutic drug utilization system by July 1, 1989. The bill also required the TDUR contractor to submit quarterly reports to DHS and the Legislature, analyzing therapeutic outcomes of the pilot project and allowing comparisons to baseline Medi-Cal experience. The pilot project was implemented in Sacramento, Stanislaus, and Kern counties. The requirement to make comparisons to baseline Medi-Cal experience led to the selection of San Diego and Fresno counties to serve as a control group. The control group was to provide baseline Medi-Cal Experience.

The contractor's analysis was to minimally include reporting on the following items:

- ❑ Compatibility of medication to diagnosis
- ❑ Overprescribing of drugs
- ❑ Prescription of contraindicated or incompatible drugs
- ❑ Number of days of drug-induced institutionalization, i.e., Medi-Cal hospital, skilled nursing or intermediate care facility days, resulting from inappropriate drug therapies
- ❑ Number of days of drug-induced institutionalization avoided due to changes made in drug regimens as a result of pilot project drug therapy standards and operations.

These reports were to be prepared once at the beginning of the pilot project (in July, 1989), quarterly, and at the conclusion of the pilot project.

Finally, AB 2606 provided that the Auditor General shall provide or select an appropriate contractor to complete an evaluation of the TDUR system with respect to:

- Impact on institutionalization of Medi-Cal eligibles as a result of the TDUR review process
- Cost impact of the TDUR process
- Preparation of an evaluation report, for submission to DHS and the Legislature by no later than May 1, 1991, with recommendations regarding TDUR cost-effectiveness and continuation.

This independent TDUR Evaluation Report has been prepared by Ernst & Young for the Auditor General to meet the applicable objectives of AB 2606.

Overview of TDUR Pilot Program

Utilization Review (UR) is a process which has been in existence at least since the issuance of the Medicaid General System Design in 1971. It is a means for completing retrospective evaluation of medical services. When assisted by computerized data processing it is a powerful tool for monitoring provider and recipient utilization patterns and exercising control over Medicaid fraud and abuse.

As a component of UR, Drug Utilization Review (DUR) focuses on the requirements for monitoring and controlling prescription drug services. In addition to controlling fraud and abuse of prescription drug services, DUR can set for itself the objective of improving the health of program eligibles through improved drug therapies. Many in the medical community believe that improving health care by focusing on patients' *therapeutic* issues can reduce the incidence of institutionalization among these eligibles. As a result health care costs should decline.

AB 2606 mandated a TDUR pilot program in California. The intent of this pilot program is to determine if TDUR can have an impact on institutionalization and be cost-effective in California.

The State Department of Health Services, as mandated, is responsible for administration of the TDUR pilot program. DHS initiated a competitive procurement process to obtain contractor services for the pilot program's implementation and operation. As a result of the competitive procurement, DHS selected The Virginia Computer Company (VCC) as the TDUR contractor. The contracted amount was $1,336,848. The pilot program commenced implementation on January 1, 1989 and was fully operational by July 1, 1989. The pilot will conclude June 30, 1991.

The Master Therapeutic Drug Utilization Review Committee has responsibility for giving the pilot program guidance in policy and has ultimate authority over standards and procedures for pilot program operations. Members of this committee are appointed by the Governor. Membership includes representation from the Health and Welfare Agency, the Department of Health Services, California schools of pharmacy, licensed pharmacists and physicians, drug manufacturers, and program beneficiaries. As many as 17 members can be appointed; however 10 to 15 members have actually served on the committee throughout the course of the pilot.

Per the requirements of the DHS contract, VCC's responsibilities in the pilot program are basically four-fold:

- Perform TDUR program operations, including management of pilot county Drug Utilization Review Committees (DURC), administrative activities for the TDUR reviews including manual case tracking activities, and status reporting to DHS
- Operate and maintain the TDUR computer processing system
- Perform pilot program evaluation activities, including production of quarterly statistical reports and two special summary evaluation reports
- Report results quarterly to the Master Therapeutic Drug Utilization Review Committee.

Once a month throughout the duration of the pilot program, VCC's automated exception processing system reviews Medi-Cal fee for service claims provided by DHS. Recipient medical history profiles are generated for those Medi-Cal patients meeting preset exception criteria. This exception criteria represents a set of tests applied to Medi-Cal claims history tapes. Medi-Cal eligibles whose record of medical services meets pre-defined parameters for abnormal, or dangerous, usage of drugs are flagged as exceptions.

Within the TDUR computer system, an encoded master file contains specific exception criteria which are used in selecting the monthly recipient exceptions. Exceptions are of two types - qualitative and quantitative. In general, definitions of therapeutic criteria used to select the exception patients are based upon the following types of data:

- ❑ Patient diagnosis
- ❑ Patient age and sex
- ❑ Prescribing characteristics related to concurrent use of potentially inappropriate drugs
- ❑ Pre-determined levels of acceptable prescription activity.

The therapeutic criteria logic was defined by VCC medical professionals. The profiles of "problem" patients, generated automatically on a monthly basis, are the primary data source used in ongoing TDUR operations. The sequence of monthly activities performed in the pilot program is as follows:

- ❑ Medi-Cal claims history files are reviewed by computerized exception criteria and patient profiles are generated
- ❑ The patient profiles are reviewed by DUR committees comprised of doctors and pharmacists who select a portion for follow-up notification
- ❑ Letters are mailed to the care providers of the above selected patients to notify them of a potential problem with their patient
- ❑ Follow-up profiles are generated automatically for all patients selected for provider notification
- ❑ The follow-up profiles are reviewed by the DUR committee members to determine if the problems were corrected
- ❑ The profiles are marked to indicate whether or not the problem was corrected and, if deemed appropriate, the case is closed. In some instances, cases remain open pending the resolution of special actions taken, such as referral to DHS.

Pilot program evaluation activities are segregated from ongoing TDUR operations. VCC initially subcontracted its responsibilities for pilot program evaluation to the Consolidated Consulting Group (CCG), including the production of statistical quarterly reports to be submitted to the Legislature. Unfortunately, CCG went out of business early

in the project. There was considerable delay in replacing CCG, and it was not until approximately September of 1990 that Stanford Research Institute (SRI) was employed to perform a summary evaluation. To date, only two quarterly reports have been submitted by VCC/SRI. These two reports do not provide the evaluative data required by AB 2606, Section 4. These reports were to include an analysis of TDUR therapeutic outcomes in a manner that would allow comparisons with the experience analyzed for the period of July 1, 1987, until the implementation of the open formulary TDUR pilot program. DHS has indicated that the deviations from planned quarterly reporting are a result of:

- Delays experienced in procuring a subcontractor to fulfill the program evaluation responsibilities
- Problems with the original study design
- Lack of reliable data being produced by the TDUR operations component.

A more detailed discussion of the theory and description of TDUR is included in Appendix C.

Evaluation Objectives and Scope

The overall goal of this evaluation, as set forth in AB 2606, was to determine the cost-effectiveness of the TDUR program and to develop recommendations regarding its continuation. Specific evaluation objectives related to this goal were as follows:

- *To evaluate the impact on institutionalization of Medi-Cal eligibles by operation of the Medi-Cal therapeutic drug utilization review process*
- *To evaluate the cost impact of the Medi-Cal therapeutic drug utilization process*
- *To determine whether the TDUR system is cost-effective*
- *To recommend whether the TDUR system should be continued.*

Subsequent to the approval of this legislation, federal legislation was passed requiring a Drug Utilization Review Program to be implemented by all states on a statewide basis. Therefore, evaluating the merits of continuing the program is no longer an important stand-alone objective in terms of legislative decision-making. What remains important,

however, is the potential characteristics of a California DUR Program. To the extent that this evaluation study can help structure an efficient statewide program, the completion of this evaluation objective remains important.

Summary of Evaluation Methodology

The contractor selected by DHS to operate TDUR was required to compare results of the open formulary TDUR against baseline Medi-Cal experience drawn from the Closed Formulary/Prior Authorization system in use in the remainder of the state. Because of the denial of the Department's request for waiver to implement an open formulary pilot, this was not done. AB2606 provided for the implementation of TDUR independent of the open formulary should the federal government deny the Department's request for waiver. It was determined by DHS and their contractor, VCC, to use the closed formulary/prior authorization system in the pilot counties and compare the result of the pilot to a control group composed of selected non-pilot counties.

The Auditor General's audit was to have reviewed the reports prepared from this comparison. This comparison has not yet been performed. As a result we designed a methodology to gather sufficient data to respond to the requirements of AB 2606 in completing an independent review of the open formulary TDUR pilot program.

Our methodology relies on an independent review of patient profiles produced by the computerized exception criteria, interviews with local Drug Utilization Review Committee (DURC) members, and interviews with employees of the Department, VCC and their subcontractor, SRI. Our findings, conclusions and recommendations are based upon the data gathered from each of these sources.

Our independent review of patient profiles resulted in categorizing the patients into two groups. The first group, which we call the Response Group, includes those individuals whose therapeutic care improved after TDUR intervention. The second group is called the Non-Response Group, and includes those patients whose medical profiles did not show improvement in drug therapies. Our analysis considered the differences in

service utilization (incidences of Medi-Cal claims) between these two groups. We used service utilization rates to determine the impact and cost-effectiveness of the program. Our methodology is described in detail in Appendix D, including the definition of terms which are used in our Audit Results. A summary of the rationale for our approach is also included in Appendix D.

AUDIT RESULTS

ERNST & YOUNG

AUDIT RESULTS

I.

THE CALIFORNIA TDUR PROGRAM REDUCED UTILIZATION OF PHARMACY, OUTPATIENT, AND INPATIENT SERVICES FOR MEDI-CAL ELIGIBLES IN THE PILOT COUNTIES

During the period from July 1, 1989 through August 30, 1990 there was an overall reduction in medical service utilization for Response Group eligibles included in our analysis. Response Group eligibles are those Medi-Cal eligibles in the pilot counties whose exception profiles resulted in program intervention, and whose re-review profiles reflect a decrease in Medi-Cal services and costs over the original review period. Results from our review of individual case circumstances indicate this decrease was likely an effect of the TDUR intervention. We measured service utilization using the average number of claims per eligible for pharmacy (drugs and other prescription items), outpatient (doctor, emergency room and clinic visits), and inpatient (hospital and long-term care days) services. In our analysis we used a comparison group called the Non-Response Group. Non-Response Group eligibles are those Medi-Cal eligibles in the pilot counties whose exception profiles resulted in program intervention, but whose re-review profiles reflect no change in Medi-Cal service utilization over the original review period. We found three significant outcomes regarding service utilization. These are discussed below.

1. Both Response and Non-Response Groups Utilized Less Pharmacy and Outpatient Services After TDUR Contact; However, The Response Group Experienced a Greater Reduction Than the Non-Response Group

Drug utilization decreased by 25 percent for the Response Group after providers were contacted and 18 percent for the Non-Response Group. Outpatient service utilization showed an even greater decrease in both groups, 34 percent and 21 percent, respectively. Table 1, on the following page, shows the number of Medi-Cal claims for each group, before and after contact. Table 2, following Table 1, shows the number of *recipients* who submitted the Claims in Table 1 (i.e., multiple Claims per recipient), both before and after TDUR contact.

Table 1

Number of Medi-Cal Claims Before and After TDUR Contact

Service Category	Before TDUR		After TDUR		Percentage Change	
	Response Group	Non-Response Group	Response Group	Non-Response Group	Response Group	Non-Response Group
Pharmacy (Services)	6,806	4,932	5,105	4,034	(25%)	(18%)
Outpatient (Visits)	1,422	1,068	931	841	(35%)	(21%)
Inpatient-Hospital (Days)	91	57	55	81	(40%)	42%
Inpatient-Long-Term Care (Days)	688	743	813	742	18%	(.1%)

Source: Ernst & Young review of VCC patient profiles produced for services from July 1, 1989 through August 30, 1990.

Table 2

Number of Recipients Submitting Medi-Cal Claims Before and After TDUR Contact

Service Category	Before TDUR		After TDUR		Percentage Change	
	Response Group	Non-Response Group	Response Group	Non-Response Group	Response Group	Non-Response Group
Pharmacy	391	360	383	345	(2%)	(4%)
Outpatient	236	192	214	169	(9%)	(12%)
Inpatient-Hospital	18	15	10	13	(44%)	(13%)
Inpatient-Long-Term Care	12	13	17	14	42%	8%

Source: Ernst & Young review of VCC patient profiles produced for services from July 1, 1989 through August 30, 1990.

Table 3, on the following page, shows the average number of services per eligible in the Response and Non-Response Groups, before and after TDUR contact. These averages are calculated using the data presented in Tables 1 and 2.

Table 3

Average Number of Claims Per Recipient Before and After TDUR Contact

Service Category	Before TDUR		After TDUR		Percentage Change	
	Response Group	Non-Response Group	Response Group	Non-Response Group	Response Group	Non-Response Group
Pharmacy	17.4	13.7	13.3	11.7	(24%)	(15%)
Outpatient	6.0	5.6	4.4	5.0	(27%)	(11%)
Inpatient-Hospital*	5.1	3.8	5.5	6.2	8%	63%
Inpatient-Long-Term Care*	57.3	57.2	47.8	53.0	(17%)	(7%)

* Average Number of Days

Source: Ernst & Young review of VCC patient profiles produced for services from July 1, 1989 through August 30, 1990.

2. *Response Group Eligibles Experienced a Decrease in Hospital Utilization*

Inpatient hospital days decreased by 40 percent for the Response Group, while they increased 42 percent for the Non-Response Group. However, there is a strong likelihood that not all of the hospitalizations were a result of a problem in drug therapy. For example, several patients in the study were apparently addicted to various drugs, such as codeine, and were hospitalized for reasons probably associated with poor health. Since the number of patients hospitalized during the study was low, the inclusion of drug addicts may have affected hospital utilization rates. After careful analysis of the hospital stays for all the patients in our study groups, it is our conclusion that the overall trend toward lower hospitalization rates for Response Group patients (other than drug addicts) is valid.

3. *No Conclusions Can Be Drawn About the Effect of TDUR on Eligibles in Long-Term Care Facilities*

The data showed an increase in long-term care days for Response Group eligibles and no change in the Non-Response Group. However, we determined that this unusual outcome could not be attributed to the effects of TDUR. This determination was based on a number of factors. First, the amount of elapsed time between the initial profile and the re-review profile was insufficient to accurately reflect change which could be attributed to TDUR contact. From the SRI Quarterly Report of December 1990 we know that the average length of stay for a long-term care episode is over 280 days. The likelihood of seeing original admissions in a ninety-day claims history was very small. Another problem was that, like hospitalization, the number of eligibles in the study utilizing long-term care

services was low. Finally, it was reported to us by several DURC members that long-term care eligibles were generally ignored in the contact process. This is because DHS, independent of TDUR, exercises monthly utilization review of the drug therapies of long-term care residents. Although some provider notifications were sent with respect to long-term care eligibles, we determined that this group was reviewed differently by different committees, and that this inconsistency may have contributed to the unusual outcome observed from our analysis.

CONCLUSION

VCC and the Drug Utilization Review Committees determined that at least 400, and possibly 720, eligibles improved with respect to drug therapies as a result of TDUR contact during the pilot. Our analysis of the medical histories of 800 eligibles contacted showed that this improvement did occur, that the improvement was likely a result of TDUR, and that those who improved reduced Medi-Cal service utilization. Based on the reductions of utilization of pharmacy, outpatient and hospital inpatient services, we concluded that the TDUR program had a positive impact on those eligibles contacted.

II.
BASED ON MEASURED DIRECT SAVINGS, THE PILOT TDUR PROGRAM WAS NOT COST-BENEFICIAL

Direct cost benefit was only $123,591, while program costs were $1,722,000 in the review period. Our review of patient profiles showed that VCC and DURC members were conservative in their estimate that 400 eligibles improved during the pilot. We determined that there were potentially a total of 720 eligibles who may have improved as a result of TDUR contact. The additional 320 eligibles represent a potential eighty percent increase in the impact attributable to the program. For this reason, our estimation of the cost impact of TDUR was calculated based on a range, with the low end representing the measurable impact on the 400 eligibles, and the high end reflecting an additional eighty percent cost savings. We consider the latter dollar figure to be the maximum, direct cost-benefit achieved by this program during the fourteen-month sampling period.

Using average service costs, derived from SRI data, of $16 per pharmacy claim, $24 per outpatient claim, and $750 per hospital day, the range of direct benefits attributable to TDUR was from $66,000 to $123,591. *Long-term care costs were ignored in this analysis because long-term care outcomes were determined to be inconclusive.*

The calculated benefit, on the low end, reflects savings from reduced utilization in the Response Group, assuming that the entire reduction is attributable to TDUR. This is unlikely, since reductions of a smaller nature occurred in the Non-Response Group for both pharmacy services and outpatient services (but not inpatient hospital days), as shown previously in Table 1. In general, we cannot attribute the Non-Response Group reductions to TDUR.

Table 4, on the next page, presents a breakdown by service category (excluding long-term care) of the measured cost-savings from our review. Cost-savings are shown for the Response Group alone, for the Response Group less the Non-Response Group, and for the Maximum Direct Benefit attributable to the program. Total cost-savings for the Response Group less the Non-Response Group are actually higher than total savings for the Response Group alone. This is because the Non-Response Group showed an increase (versus a decrease) in hospitalization costs.

Table 4

TDUR Cost-Savings By Service Category

Service Category	Average Service Cost	Service Utilization	Response Group	Response - Non-Response	Maximum Direct Benefit
Pharmacy	$ 16	1,701	$ 27,216	$ 22,263	$ 40,073
Outpatient	24	491	11,784	9,274	16,693
Inpatient - Hospital Days	750	36	27,000	37,125	66,825
Total Savings			$ 66,000	$ 68,662	$123,591

Source: Ernst & Young review of VCC patient profiles produced for services from July 1, 1989 through August 30, 1990; SRI Quarterly Report for December 1990.

Based on measured direct savings, the TDUR pilot was not cost-beneficial. Pilot program expenses, including contractor services of approximately $719,550 and state administrative costs of approximately $1,002,450, were a total of $1,722,000 for the fourteen months covered by our review. Since the program budget did not separate one-time costs from ongoing costs, it was not possible to determine what portion of this amount may have included start-up costs. The stated costs exceeded the maximum direct benefit by $1,598,409. For the pilot to be cost-effective, indirect benefits would at least have to cover the portion of this amount which represents ongoing expenses for fourteen months.

CONCLUSION

The direct cost-savings of the TDUR pilot were insufficient to offset the cost of operating the program. If implemented on a statewide basis, the TDUR process used during the pilot would not be cost-beneficial for California unless potential indirect benefits are considerably larger than direct benefits. Costs per eligible for DUR programs in other states vary widely, from $.08 to $2.00 annually. With three million Medi-Cal eligibles, costs for California could range from $240,000 to $6 million a year. To be cost-effective, California's program would have to achieve direct and indirect savings which are at a minimum equivalent to its operating budget. If the pilot had focused on therapeutic issues as planned, the cost-savings from reduced incidences of hospitalization would most likely

have been greater. If so, it is reasonable to expect a substantial increase in overall savings would have been achieved, since the cost for each incidence of inpatient care far exceeds the cost per claim for drugs and outpatient services. To illustrate, the cost-savings from 47 pharmacy claims is equivalent to avoiding one day of hospital care for a Medi-Cal recipient.

III.

IMPLEMENTING A COST-EFFECTIVE STATEWIDE TDUR PROGRAM IN CALIFORNIA WILL REQUIRE SIGNIFICANT BUT POTENTIALLY FEASIBLE CHANGES FROM THE PILOT PROGRAM

Our final evaluation objective from AB 2606 was to recommend whether the TDUR program should be continued in California. Subsequent *federal* legislation required that a Drug Utilization Review Program be implemented by all states. Since California is now required to implement a DUR program, our response to this question focuses on *how* to implement the program versus *whether* to implement the program. It should be noted that the statewide program need not be implemented in a manner similar to that of the TDUR pilot.

Although we were unable to prove the cost-effectiveness of the TDUR pilot, we found many opportunities to improve the program for statewide implementation. The program needs restructuring to focus on improving the benefits and operating within a targeted budget. The process used during the pilot was a broad based approach, which we believe attempted to do too much for too many. With more targeted criteria and screening, the program could do a better job of alleviating drug utilization problems, and this may result in lower health care expenses for Medi-Cal recipients. Below we discuss our specific findings and recommendations regarding changes in program practices which should be considered before statewide expansion.

1. The Exceptions Criteria Used During the Pilot Program Were Not Well Focused

VCC reported that out of 17,000 exceptions flagged during our review period, only 2,000 resulted in TDUR intervention. Of those, only 400 demonstrated changes in therapy as a result of contact. This indicates that at least 88 percent of the exceptions flagged by the computer system were not appropriate for this process, and perhaps more. Members of the DURC confirmed that a significant amount of the data produced as exceptions were of little or no merit for the purposes of the TDUR program. For example, anti-coagulents represent a category for review which could yield significant cost-avoidances in hospital stays. This category was not one of those selected for review during the pilot. Also, we found other examples of poorly focused criteria: the inclusion of maximum services criteria served to distract many members from the primary purpose of the program, which is avoidance of institutionalization; members found they were unable to affect emergency room services through the intervention process; some

criteria focused on oral contraceptives, while use of this drug does not generally result in institutionalization.

The inability of the system to focus reliably on potential problems resulted in significant wasted time, reducing the benefit achieved from the program. The maximum services criteria, in particular, resulted in cases inappropriate for TDUR.

We found in our review of the MTDUR Committee meeting minutes that there was little discussion of appropriate therapeutic criteria. As we understand it, VCC implemented a standard set of criteria in use in other contracts. Little consideration was given to modify these criteria to better focus them or make them more clinically relevant to the requirements of the legislation.

2. *Screening of Medi-Cal Recipients was Too Generic*

The TDUR process attempted to screen all Medi-Cal eligibles in the pilot counties. This is unproductive, since most problems in drug therapy affect only certain populations among Medi-Cal recipients, such as the elderly. Eligibles comprising most exceptions generated during the pilot were under age 65.

3. *The Review Committees Lacked The Ability to Satisfactorily Follow-Up on Cases Selected For Intervention*

DURC members reported that their contact with providers was limited to the "form" letter sent by VCC. This method of intervention meant that the reviewers had little input into the drug therapies of recipients. They believed that the intervention process would have proven more effective and its acceptance would have been greater if the contact was by phone or some other personal means. Personal communication with providers would also have enabled reviewers to clarify key information about the patient's drug therapy to save time in the review process.

4. *Reviewers Could Have Handled Cases More Efficiently if Lab Test Data Were Available on the Patient Profiles*

The patient profiles produced by the computer system did not include information about laboratory services which may have been ordered by the prescribing physician. If this information was included, the reviewer could determine if a physician was aware of the potential for a drug reaction problem. Lab test orders on the medical profiles would thus save time in unnecessary review, letter generation, and case tracking.

5. *There Was No Feedback Loop in the DUR Process to Implement Suggestions From Those Involved in the Program*

Communication between the local DURCs and the Master Therapeutic Drug Utilization Review Committee (MTDURC) was indirect, i.e., through the contractor, VCC. While this indirect relationship appears to be consistent with the Department's original plan, it also made it very difficult to use local DURC experience to improve some of the problems discussed in this

report. For example, local reviewers were aware of problems with the exceptions criteria and ways to improve them. They were also aware of the potential benefits of establishing personal contact with providers. There were no procedures in the TDUR process to make MTDURC members aware of these potential improvements or to implement the changes if deemed appropriate.

6. The Program May Have Operated Better If Contract Administration Had Been More Rigorous

The administration of the pilot program contract did not incorporate certain practices which the state typically applies in these situations, particularly with large-dollar contracts. For example, specific milestones and deliverables were either not scheduled or not enforced; general due dates were missed frequently without penalty; quality assurance guidelines/expectations were not established; and original plans were altered without an apparent full understanding of program and/or evaluation implications. More traditional administration of future contracts should enhance overall program performance.

7. More Providers Could Have Been Contacted If the State's Provider File Was Up to Date

Although it was not possible to determine the impact of the letters which remained unsent or undelivered in this project, it seems reasonable that the more providers contacted, the greater the potential cost impact.

CONCLUSION

We believe that the potential exists for California to run a cost-effective TDUR program. However, this will require a number of changes with respect to program administration and operations. Our recommendations regarding these changes are summarized here. If implemented, they should result in greater cost savings and lower administrative costs for the state.

RECOMMENDATIONS

Our recommendations assume that the state will implement a TDUR program similar to the pilot. However, we recommend that a cost analysis of implementation alternatives be undertaken before proceeding with statewide implementation. The implementation alternatives considered should include at a minimum the use of in-house specialized state employees (i.e., pharmacists) versus DURCs (whose members probably will require compensation). Factors related to prospective versus retrospective TDUR

approaches also need to be evaluated. Finally, the state should consider the following recommendations resulting from our study:

- New exceptions criteria should be developed for the TDUR process. This improvement probably offers the greatest potential for increased savings. Changes should include eliminating the maximum services criteria and any other criteria that is not applicable to known therapeutic problems. The criteria should be diagnosis driven and screening should be limited to drugs with a high toxicity.

- Exceptions criteria should be applied more selectively. Screening should not include all Medi-Cal recipients in the state. Screening is more cost-beneficial if done on a sampling basis, particularly if it is limited to those recipients most likely to benefit. The elderly are particularly susceptible to drug therapy problems, but they were not targeted in the pilot.

  It is the intent of the first two recommendations to assist the state with developing an approach similar to Medi-Cal's fraud detection unit, which only attempts to find those cases which will result in the greatest dollar savings. For a TDUR program, this means focusing on reducing hospitalization and long-term care utilization.

- The Department of Health Services should provide more rigorous administration of contractor conduct. This would include developing written procedures and ensuring that they are followed, setting specific due dates and quality assurance guidelines for deliverables, and enforcing penalties when they are not met.

- The Master TDUR Committee members should develop relationships with county review committees and elicit feedback from them regarding program operations. When appropriate, suggestions should be incorporated into the policies and guidelines for the state's TDUR process. There should be a mechanism for ensuring that the policies and guidelines (new and old) are implemented consistently throughout the state.

- Local review committees, if used, should be given the ability to follow up on cases selected for intervention, if needed. Improving the criteria and reducing the number of individuals screened each month will enable closer monitoring of problem cases, but procedures must be implemented to support this activity.

- Lab test data should be included as part of the information available to local review committees. This will help eliminate unnecessary interventions and facilitate the review process. However, a safeguard should be developed to ensure that physicians treating Medi-Cal patients do not adopt a policy of requesting unnecessary lab tests as a defensive reaction to TDUR interventions. This type of backlash would likely cause a net increase in Medi-Cal expenditures rather than a reduction of costs.

- The accuracy and completeness of the state's provider file should be improved. This will increase the number of providers who can be contacted when interventions are applied.

APPENDIX A

RESPONSE TO AUDIT

ERNST & YOUNG

DEPARTMENT OF HEALTH SERVICES
714/744 P STREET
P.O. BOX 942732
SACRAMENTO, CA 94234-7320
(916) 445-1248

April 19, 1991

Mr. Kurt R. Sjoberg
Auditor General (Acting)
Office of the Auditor General
660 J Street, Suite 300
Sacramento, CA 95814

Dear Mr. Sjoberg:

Thank you for the opportunity to respond to the draft audit report on the "Cost Effectiveness of the Medi-Cal Therapeutic Drug Utilization Review Program". The staff developing the report were required to assimilate a significant amount of information about a complex program area in a short amount of time. The report reflects that they achieved a good understanding of the program.

The Department cannot concur with the findings set forth in the report, and we strongly recommend that any conclusions about the cost effectiveness of retrospective Therapeutic Drug Utilization Review (TDUR) be based on the research effort being conducted by SRI International, as part of the legislative mandate for the pilot operation of TDUR. The recommendations made in the report for improving statewide TDUR operation, however, could enhance the value of this system, and as the Department moves to implementation of a statewide TDUR program they will be considered. (1) *

The Department has repeatedly expressed its concerns about the methodology used in determining the cost effectiveness of the TDUR project by the Auditor General study team. Our concerns were detailed in a February 6, 1991 letter to Mr. William S. Aldrich of your staff, a copy of which is enclosed. The significant concerns are summarized here. First, since the study design included only a small number of participants from the TDUR treatment group, any conclusions about the effectiveness of TDUR cannot be made. This methodology does not include a true control group; there is no identification of changes induced by outside influences which could affect individual outcomes; and there is no assurance of comparability for eligibility or service availability. (2)

Second, upon review of the data used in assessing the cost effectiveness of TDUR on this population, no account was made for the effect the lag in receipt and processing of claims for services has on cost considerations. Only three months of payment data was used in determining cost impact. This represents, for example, only about 20 percent of county hospital claims and 47 percent of community hospital claims. Excluding these components significantly misrepresents the costs both before and after TDUR intervention, and renders any conclusions reached about TDUR cost effectiveness meaningless. (3)

*Circled numeric references to Ernst & Young's response to the Department's response, included in Appendix B.

The report recognizes that the Department is required to implement a statewide TDUR program in response to federal requirements included in the federal Omnibus Budget Reconciliation Act of 1990 (OBRA '90). The Department is just beginning to develop the requirements for this project. The recommendations made by the Auditor General study team will be considered in our development of TDUR specifications, and incorporated as appropriate in our contract requirements.

Sincerely,

Tamera Mitchell

for Kenneth W. Kizer, M.D., M.P.H.
Director

Enclosure

DEPARTMENT OF HEALTH SERVICES
714/744 P STREET
P.O. BOX 942732
SACRAMENTO, CA 94234-7320
(916)427-7760

Mr. William S. Aldrich
Deputy Auditor General
Office of the Auditor General
State of California
660 J Street, Suite 300
Sacramento, CA 95814

Dear Mr. Aldrich:

Department of Health Services' staff met on January 23rd with you and Ernst & Young staff. A subsequent meeting was held with Ernst & Young staff on January 28th. At those meetings Ernst & Young explained their strategy for fulfilling the Auditor General's obligations under the legislation enacted to establish the DUR project.

DHS staff were surprised that Ernst & Young is proposing to independently undertake a research project, rather than to review the efficacy of the conduct of the DUR pilot project. We are concerned in that Ernst & Young is proposing to undertake research on exceptions-based measures. Stanford Research Institute International (SRI), the professional research firm evaluating DUR, has advised that this is not an area for fruitful investigation. SRI is one of the top research organizations in the country staffed by highly qualified professionals with actual experience and advanced academic degrees in evaluation research, statistics, and health care related disciplines. Our staff concur with SRI's findings in this area.

From our lengthy meeting we learned that the Ernst & Young proposal is a relatively simple plan to evaluate this complex project. They want to examine the approximately 400 exceptions profiles identified to date by Mr. Glen Spaulding, the project's DUR process manager, as having positive outcomes from DUR intervention. Exception profiles are 90-day paid claims histories on individuals who have been targeted as having received inappropriately prescribed drugs. The prescribers involved are notified thereof. Positive outcomes were not clearly defined by the subcontractor in the meeting. Our questioning revealed that neither Mr. Spaulding nor Ernst & Young have developed objective, written criteria for selecting exceptions showing positive change. Instead, Ernst & Young is relying solely on the subjective views of the project's DUR process manager.

Ernst & Young will have the profiles from Mr. Spaulding reviewed by their own professional staff, comprised of an out-of-state physician and pharmacist, to see if they agree with him. Then Ernst & Young proposes to take an additional 400 exception profiles from the DUR test group and in some manner compare the health service costs for that group with Mr. Spaulding's group. The average costs per service for all Medi-Cal eligibles would be used for making that comparison.

Departmental staff pointed out the problems with this approach in both the January 23rd and January 28th meetings. We brought up all the following problems:

1. SRI has clearly stated that exceptions data are not useful for determining whether DUR is or is not cost beneficial to the State;

2. Ernst & Young is not attempting to create a double blind study of the exceptions. Instead, they are just accepting the subjective conclusions of the project's DUR process manager;

3. Ernst & Young has not developed explicit, objective, written criteria for classifying exceptions as showing positive changes occurring specifically from DUR intervention;

4. The 400 comparison profiles are also drawn from the DUR test population instead of the actual control population created for the DUR pilot project.

Based on the above, we do not feel the study meets the criteria of having a scientifically sound research methodology. Consequently, we are perplexed about the significance of any findings to be produced by Ernst & Young. Are they going to state that the results apply only to the few profiles they reviewed or are they going to extrapolate their results to a larger population? If they do so, what is the larger population to which they would generalize their findings?

We feel that Ernst & Young could not generalize any of their findings to a large population due to the faultiness of their research design.

I would prefer to see the Auditor General's Office review the Department's process of implementing and evaluating the DUR pilot project, rather than independently undertaking a questionable research project. If your Office decides to pursue this latter course of action after considering our objections, we recommend several tasks, as a bare minimum.

1. Ernst & Young give all the approximately 800 (400 "positive change" and 400 "no change") exceptions to their pharmacist and physician for evaluation on a <u>double</u> <u>blind</u> basis;

2. The 400 comparison exceptions should be selected randomly from the population receiving letters but which Mr. Spaulding was unable to find any change;

3. Ernst & Young develop explicit, objective criteria for classifying exceptions as showing positive changes vs. no changes as a result of DUR intervention;

4. Ernst & Young not use "average costs per service" for all Medi-Cal eligibles.

Regarding item 4., the majority of DUR exceptions are the elderly and disabled who often incur health care costs which would not be paid by Medi-Cal. It would be easy for DHS to retrieve from our Medi-Cal paid claims the cost and service data needed by Ernst & Young.

We would access our computerized Medi-Cal paid claims files so Ernst & Young does not have to manually compile and key enter the data they need. If Ernst & Young is concerned about the accuracy of our data, they could randomly spot check and compare their recipient profiles data with the retrieved paid claims data.

In closing, I want you to know that we will assist, if needed. However, our assistance should not in any manner be construed as supporting or approving this project and its findings. We urge you to abandon the quasi-research. We believe it will only lend confusion to the real research findings from the DUR project.

Please be aware that the research being undertaken by SRI was not devised in a vacuum nor spelled out by anyone with a particular point of view. The whole project has been reviewed and approved by a number of groups, including The Virginia Computer Company, the Department of Health Services, and the Medi-Cal Therapeutic Drug Utilization Review Committee (MTDURC). The MTDURC oversees the entire DUR project. It is comprised of individuals appointed by the Legislature and the Governor's Office. We believe that the research outcomes are duly satisfied by this undertaking.

If you have question, please contact me at 427-7760.

Sincerely,

James P. Parks, Chief
Office of Health Systems Financing

cc: See attachment

cc: John Keith 8/1476
Jean Crew 8/1476
Kelly Klemin 9/1125
Milt Kuschnereit 8/1340
Lena Terra 8/1340

APPENDIX B

ERNST & YOUNG COMMENTS ON THE RESPONSE

ERNST & YOUNG

ERNST & YOUNG'S COMMENTS ON THE RESPONSE

This appendix includes our response to the Department's response to the audit report which is included in Appendix A. We have organized our response to address the issues raised by the Department, in the order raised. Before doing so, the following clarification is made.

In their response, the Department references its letter to the Office of the Auditor General (OAG) of February 6, 1991. The Department fails to mention the Auditor General's response to the Department of February 26, 1991 in which the Department's concerns are addressed. The letter submitted to DHS on February 26, 1991 is included in this appendix following our response. This letter makes note of several key points.

- First, the cost-effectiveness of the program could not be determined from existing TDUR pilot program data reports. This is because (1) the Department believes the data reports produced by VCC were inaccurate; (2) the Department did not complete its planned exceptions-level analysis; and, (3) the DHS contractor, VCC had not produced any of the quarterly analysis that were planned under contract. As a result, independent analysis was necessary to evaluate the cost-effectiveness of the program.

- Second, because of the lack of reliable exceptions-level data for the pilot and control populations in this study, and the lack of an exceptions-level analysis, there is insufficient data to statistically analyze the program's performance in a reliable manner. This necessitated an approach whereby we had to analyze the actual case profiles that were acted upon by the local DUR committees.

- Third, because of deficiencies in the data produced by the program, we found it necessary to evaluate all case profiles for those eligibles adjudged to have benefitted through improved drug regimens after TDUR contact. This was done to compensate for the deficiencies in pilot program reporting.

The remainder of this discussion addresses issues raised specifically in the response from DHS. Numbers in the left-hand margin below reference specific portions of the Department's letter in Appendix A.

(1) The Department recommends that any conclusions about the cost-effectiveness of the Therapeutic Drug Utilization Review (TDUR) pilot program be based on research being conducted by SRI International, VCC's subcontractor. A draft of this analysis was originally scheduled for publication in March 1991; however, actual delivery has been delayed until May 24, 1991, well after the issuance date of

our report. Irrespective of the delivery date, on page C-24 of our report we describe our concern with the planned deviation in the evaluation approach taken by the Department and SRI. This includes abandoning exceptions-level measures in evaluating the cost-effectiveness of the program. By so doing, an important control planned in the program's original study approach was eliminated. As a result, the SRI analysis will be unable to confirm that any observed differences between the pilot counties and the control counties resulted from the program. Consideration of this change played a significant role in developing our own methodology to assess cost-effectiveness of the program by examining exceptions-level data in the form of case profiles.

(2) The Department expressed concern over our methodology in its letter of February 6, 1991. The Auditor General and Ernst & Young responded on February 26, 1991. We noted then that our methodology was modified where feasible to address some of the Department's concerns. Our letter, included after this discussion, also addressed the issue of the use of a small number of participants. This issue is also addressed in our report on pages 13 and 14, and further on pages D-1 through D-3. While small, the approximately 400 eligibles in the Response Group represent 100 percent of the eligibles found by VCC and the local DUR committees to have improved under this program following intervention. Any other beneficiaries of this program would have to result from indirect causes (e.g., provider education), and from direct causes. We state on pages S-1 and D-5 of our report that we did not consider indirect benefits in our evaluation. We also state on page S-1 that indirect benefits would have to be on the order of 13 to 24 times the direct benefits to make the program cost-effective.

Our comparison group was selected for the reasons described in our letter, which follows this discussion. Had the Department and its contractor produced and maintained accurate exceptions-level data, or produced exceptions profiles for eligibles in the control counties, we could have used a control group which was more isolated from potential bias. Unfortunately, the necessary data was not produced during the program.

Additionally, our methodology was designed to compensate--as best feasible under the circumstances--for the absence of external variables. This was accomplished by performing an independent physician review of each case judged by VCC and the local DUR committees to have improved after intervention, and of an additional review of sample cases from the Contact Group which did not reflect improvement. This evaluation was performed in a double-blind manner.

(3) Our report states on page 9 that our evaluative period includes the fourteen months from July 1, 1989 through August 30, 1990. We did not include data from the full two years of the pilot program because the program is not yet complete and claims do experience a lag time in billing. Our sample period was selected precisely because of billing lags experienced in Medi-Cal claims. We also point out that we did not consider program costs for more than the fourteen month evaluative period.

While additional benefits are conceivable if the evaluative period is extended, the proportionately larger cost of the program must also be considered. The full program cost is $2,952,000; this represents approximately twice the program costs considered in our evaluation.

Telephone:
(916) 445-0255

STATE OF CALIFORNIA
Office of the Auditor General
660 J STREET, SUITE 300
SACRAMENTO, CA 95814

Kurt R. Sjoberg
Auditor General (acting)

February 26, 1991 C-775

James P. Parks, Chief
Office of Health Systems Financing
Department of Health Services
714/744 P Street
Sacramento, California 94234-7320

Dear Mr. Parks:

This letter responds to your letter of February 6, 1991, regarding the Ernst & Young methodology for evaluating the Therapeutic Drug Utilization Review (TDUR) pilot program. In the attached comments, Ernst & Young provides a detailed response to each of the issues you raised. The Ernst & Young comments indicate that, in the areas where it was feasible, they had already modified their technical approach based on meetings with your staff. In other areas, Ernst & Young provide detailed explanations as to why it is not feasible or practical to implement some of your suggestions.

The statutes that established the TDUR pilot program requires the Department of Health Services (DHS) to submit quarterly reports to the Legislature, analyzing TDUR outcomes, starting in September 1989. These statutes also require the auditor general to submit a report to the Legislature by May 1, 1991. Due to subcontractor problems, the department has not fully complied with the requirement to analyze TDUR outcomes and as a result there is insufficient data to assess program effect. In view of this we believe Ernst & Young's methodology of analyzing the exceptions provide a reasonable and practical approach to evaluating and providing some information on program results by May 1, 1991.

Thanks to you and your staff for your cooperation and help in refining our methodology. As usual, a draft copy of Ernst & Young's report will be sent to the agency secretary for your review and comment prior to public release. If you have any questions, please do not hesitate to contact me.

Sincerely,

WILLIAM S. ALDRICH
Deputy Auditor General

cc: John Rodriquez, Deputy Director
Department of Health Services

Milt Kuschnereit
Department of Health Services

Perry Spiropoulos, Project Manager
Ernst and Young

Attachment--Ernst & Young Letter

Ernst & Young

555 Capitol Mall, Suite 650
Sacramento, California 95814
Telephone: (916) 443-6756
Fax: (916) 443-0774

February 20, 1991

Mr. William S. Aldrich
Deputy Auditor General
Office of the Auditor General
State of California
660 J Street, Suite 300
Sacramento, California 95814

Reference: Response to the Department of Health Services (DHS) Letter to the Auditor General Regarding the Ernst & Young Methodology for Performance of the Therapeutic Drug Utilization Review (TDUR) Pilot Program Evaluation

Dear Mr. Aldrich:

As you requested, we have reviewed the above referenced letter from DHS and have prepared a reply to each of the issues raised therein. The comments and opinions expressed in this letter were not surprising considering that we have met at least twice with the DHS to discuss our methodology. Similar opinions and comments were expressed during these meetings and, where feasible, we have modified our technical approach based upon the Department's suggestions.

This reply is organized under the following headings:

- *Undertaking an Independent Analysis of TDUR*
- *Use of Exceptions-Based Measures*
- *Use of a Double-Blind Approach*
- *Development of Specific Written Criteria*
- *Comparison Profiles Not Drawn from the Control Population*
- *Use of Average Costs Per Service*
- *Summary of Key Points.*

1. Undertaking An Independent Analysis of TDUR

DHS states on page one of their letter that they were surprised that Ernst & Young was proceeding to independently undertake a research project, rather than merely "...review the efficacy of the conduct of the DUR pilot project." Under normal circumstances we might agree with the Department's preferred approach. Unfortunately, none of the products essential to carrying out this approach have been produced. For example, the Quarterly Statistical Reports to the Legislature were to contain evaluative data on a cumulative basis starting September 1989. To date, only two of these reports have been issued, and neither contains reliable data to assess the program's effect.

We also were told by DHS that a substantial part of the statistical reports generated to date by the Virginia Computer Company (VCC) contain inaccurate or questionable data. Use of this general program data for evaluative purposes is not advisable.

Finally, the Stanford Research Institute (SRI) is under contract with VCC to conduct an evaluation of the TDUR program. SRI's report is planned for delivery sometime in early April, with a draft scheduled before that date. However, no firm date for either report has been set by DHS. Our report for the Auditor General is due no later than April 5, 1991, so it can be reviewed and delivered to the State Legislature by May 1, 1991. Thus, we cannot wait for the SRI report, which may or may not be available for detailed analysis prior to the time the Ernst & Young report is due to the Auditor General.

On page 2 of their letter, DHS suggests that in lieu of an independent new TDUR analysis, we should undertake a review of "...the Department's **process** of implementing and evaluating the DUR pilot project..." (emphasis added). We disagree. Everyone involved in this project recognizes that significant problems occurred in the **process** of implementation. Original evaluation plans were scratched, data in reports were not reliable, progress reports were not issued, etc. We see no practical benefit to the Legislature in a process review of the program.

For the reasons stated above, we concluded that a separate analysis of TDUR effectiveness is essential if we are to meet our contractual obligations to the Auditor General and the Legislature.

2. Use of Exceptions-Based Measures

The primary objection to our methodology which was raised by DHS is that we are undertaking an evaluative study which uses "exceptions-based" measures. The Department does not believe this is a valid approach and states that SRI agrees that the use of exceptions-based data are not useful in determining the cost effectiveness of TDUR.

This DHS point is important and warrants detailed explanation. However, part of any differences of opinion on this subject might be traced to definitions of terms (e.g., "exception" versus "contact" group, etc.). Accordingly, we first set forth some definitional guidelines in the material that follows. Then we discuss SRI's (and, thus, the Department's) position as best we know it, as well as our reply. Then we provide a summary of the key points related to this subject.

Definitions

For purposes of our study, we are using the following definitions:

- *Eligibles* -- refers to all Medi-Cal recipients in both the pilot and control counties (about 300,000 in each)

- *Exceptions* -- refers to the number of exceptions generated by the computer based upon specific drug utilization criteria (about 17,000 in the pilot counties; not applied to the control counties)

- *Contacts* -- refers to a subset of "exceptions" (above) in the pilot area which were identified by the TDUR Committees as warranting a notification letter regarding potential inappropriate drug utilization (about 2,000 in the pilot counties)

- *Responses* -- refers to a subset of the "contact" group (above) who responded to the contact letter, or who appeared to have modified drug utilization patterns after the contact letter was issued, even though no written response was received (about 400 in the pilot counties).

It is important to note that Ernst & Young's approach uses only the "contact" and "response" groups (about 400 profiles from each group). We are not planning on generalized statistical analysis using the "exceptions" level data. We believe that the Department's use of the term "exceptions-based measures" relates specifically to the exceptions group of 17,000 (and its comparative cohort in the control group).

SRI's Position on Use of Exception-Based Measures

We also assume the DHS reference to SRI's position is based upon a project memo from the latter company to DHS dated January 21, 1991. That memo raised four problems with the use of exceptions-based measures in the program evaluation. These points are summarized below, together with our response. We assume that SRI's comments pertain to the "exceptions" group as just defined in this letter. We do not know if they would apply their comments to the "control" and/or "response" groups as well.

- *Changing Exceptions Criteria Over Time*

  SRI points out that the maximum services component of the exceptions-level criteria have changed several times during the pilot study. In turn, this means that the exceptions population would be defined differently at different times, and that this would invalidate the reliability of using the information to define a reference population.

  We agree with SRI that the continuing changes to maximum services exceptions-level criteria resulted in an exceptions population that is defined differently at different times. SRI goes on to say that because of the structure of VCC's database, claims for people with maximum services exceptions cannot be separated from those of the rest of the exceptions population.

  In our analysis, we will attempt to isolate the results of maximum services exceptions-level criteria from those of therapeutic exceptions level criteria. We will investigate whether these maximum services criteria changes had any real effect on decisions made by DUR committees to make written contact with providers, and we will investigate whether maximum services criteria resulted in any direct program benefit.

  We wish to point out that the original VCC study methodology prepared by Consolidated Consulting Group (CCG) acknowledged this problem and included a means to correct for it. Also, this situation will be true in a real-life environment; i.e., criteria will be changed as conditions warrant.

Today's medical services and Medi-Cal environment are dynamic and constantly changing. This situation cannot be ignored. It is not possible to construct a purely experimental design. CCG recognized this and offered a quasi-experimental solution which, unfortunately, was never implemented.

- *Problems Defining the Exceptions Group*

SRI discusses a number of issues which make the definition of the exceptions-level group difficult. Principally they state that the subgroup must be defined independently of the treatment (something not done in this pilot program). Not doing so would result in what SRI refers to as "paradoxical effects." They believe this represents an unavoidable problem and characterize the exceptions-level subgroups and data as unreliable for a comparative time-sensitive analysis.

We agree with SRI's assessment. On the other hand, as mentioned earlier, it is not possible to achieve a truly experimental design in today's Medi-Cal environment. The TDUR pilot was never intended to be purely experimental, and it can never be so given the dynamic environment at play in Medi-Cal.

The original proposal from VCC and CCG acknowledged this issue. The proposed analysis was characterized as quasi-experimental for just this reason. The original proposal went on to clarify the risks attributable to non-random assignment of recipients to subgroups, and to outline a complete methodology to isolate the effects of these problems from the outcome variables.

We believe that had the original methodology been followed, these problems could have been isolated and a reasonably reliable exceptions-level analysis completed. This is important, as only an exceptions or contact-level analysis can measure the **direct** effects of the TDUR program. VCC and CCG state in their *DUR Program Evaluation, Detailed Technical Proposal* of March 15, 1989 that in order to substantiate that the indirect TDUR effect was caused by the TDUR contact, it is necessary to measure the direct effect.

As VCC and CCG confirm in their document, direct effect can only be measured at the treatment (contact) level.

- *Inability to Verify or Replicate*

SRI states that they are unable to replicate the exceptions criteria used in this study. For this reason, they believe that the study results could not be verified, tested, or compared to those of other studies by the policy research community.

We believe sufficient documentation is available to understand the criteria used. The Master TDUR Committee (MTDUR), as well as the local DUR committees understood these criteria sufficiently well to make sensitive medical judgments from them. In contrast, what is not currently possible is the recreation of the computer programs which execute the data tests which comprise these criteria. We agree with SRI that replication of the outcome of the exception determination process would be very difficult, but we do not consider this a serious problem as we believe there is sufficient data to make comparisons of study results to similar studies in other states.

Also, it is our understanding that the DUR system and its documentation will become the property of the State of California at the end of the pilot program (6/30/91). We have also been told by VCC that the same DUR system is currently in operation in other states and is in the public domain.

- *Lack of Relevance to Public Policy Decisions*

SRI's letter states that few people would be interested in exceptions-based measures, "...particularly if the effects of the treatment for the eligibles-based measures are not statistically significant..."

We believe that using a subset of exceptions-level cases provides the only data available today to measure direct impact of the TDUR program on the Medi-Cal community. Without these measures, there is no way to validate indirect benefit measured at the population-level and to reliably

attribute it to the TDUR program. Therefore, our approach should have considerable relevance to public policy decision-making.

Summary of Key Points

We agree with SRI that performing an analysis at the level we propose is more complex and requires accommodation of a larger number of variables. On the other hand, as discussed in the original VCC/CCG proposal, some form of exceptions-level analysis should be completed in addition to population-level analysis for the study findings to carry any reliability as far as benefit attributable to the TDUR program.

VCC/CCG offered a methodology which addressed the concerns raised by SRI in their letter. As it turned out, the demise of CCG and the extended amount of time taken to replace that firm made it nearly impossible for SRI to complete all the tasks outlined in the VCC/CCG proposal. Also, the Department's position regarding the questionable reliability of VCC reporting contributed to a decision to abandon all but the population-level analysis. We wish to point out that problems attributable to contact or exceptions-level measures roll-up to the population-level. To ignore these issues at the subgroup level only serves to ignore the most fundamental question in this pilot: Is there any direct impact from the TDUR pilot program?

Ernst & Young has elected to proceed with an analysis of the "contact" and "response" subgroups of the "exception" group because we believe it is important to make an effort to assess the direct impact of the TDUR program on the pilot Medi-Cal community. Without a measure of direct impact, the reliability of the indirect impact is questionable.

3. Use of a Double-Blind Approach

In our previous meetings with DHS they raised concerns about our reliance on the TDUR contractor's Project Manager to identify profiles adjudged to have made some type of positive "response" in drug utilization following issuance of the contact letter. The Department suggested that when our medical/pharmacological consultants review the 800 profiles (400 each from the "contact" and "response" groups), that the profiles not be identified as to the group they were selected from. We have accepted the DHS suggestion and, accordingly, have implemented a "double-blind" evaluation of these profiles.

4. Development of Objective Written Criteria

DHS also wanted the evaluation by our specialists of the above referenced 800 profiles to be performed in accordance with specific written criteria as to what constituted a positive response in terms of modifications in drug utilization. This had been planned by our medical and pharmacological consultants, but the plan was not communicated to DHS. Consequently, we are proceeding as DHS had suggested.

5. Comparison Profiles Not Drawn from the Control Population

DHS objects to the selection of the 400 comparison profiles from the TDUR test population rather than from the control population. We agree that selection of these profiles from the control group would be preferable. However, the required comparison information was not maintained for the control group; thus, the DHS suggestion is not possible. We believe a comparison of outcomes between recipients in two groups with similar medical treatment and utilization characteristics (one in which positive change is likely attributable to the program, and on in which no change is evident) is a reasonable alternative for measuring a maximum, direct benefit attributable to the pilot effort.

6. Extrapolation of Our Evaluation Results

The DHS letter asks if we will extrapolate our results to a larger population. With respect to the pilot population of 300,000, we are evaluating the total population who directly benefitted from the pilot; this represents, in fact, only 437 recipients during the period of 7/1/89 to 8/31/90 inclusive, assuming the contractor's assessment of improvement to be accurate. Though the possibility exists for second and third order benefits resulting from provider education and word of mouth, these do not represent direct benefits and are impossible to measure at the recipient history profile level of detail.

The only extrapolation possible is from the pilot population to the State Medi-Cal population. We believe this to be a legitimate extrapolation for two reasons:

- It was the original intent of the pilot to do so. The pilot counties were selected because they appropriately represented the State as a whole.

- The exception criteria used in the pilot represent at least 98 percent of the criteria that probably would be used in a statewide program.

Our plans, therefore, call for extrapolation of the pilot results to the State's total Medi-Cal population.

7. Use of Average Costs Per Service

DHS objects to our planned use of average costs per service in performing the cost analysis of TDUR program impact. The Department points out that many of the exceptions group are the elderly and disabled whose health care costs are not paid by Medi-Cal. DHS suggests instead that they provide us with specific Medi-Cal cost data from Department files.

We appreciate the Department's offer. However, our ability to provide DHS with the specific selection criteria for this data is dependent on our completion of the data entry of review/re-review profiles into our data base. Once this is complete we can provide the Department with a list of recipients and service periods for preparation of Medi-Cal Claim Detail Reports (CDRs). These reports will provide us the service level payment information desired by the Department. Our concern, however, is that we may not be able to produce the selection criteria soon enough for DHS to return the CDRs before our analysis is finished.

Therefore, we will continue our plans to use average costs per service in performing the cost analysis of the TDUR program. Should the more detailed data become available with sufficient time to adjust our analysis, we will do so.

8. Summary of Key Points

We wish to summarize the key points included in this response to the DHS letter. These are organized under the following headings:

- Items which are no longer relevant
- Items which we agree with in principle but cannot do as a practical matter
- Elements of our approach which are consistent with or similar to the original TDUR analysis approach.

Items Which are No Longer Relevant

This category specifically addresses the methods we are using to evaluate the approximately 400 "positive change" and 400 "no change" contact exceptions. We believe the concerns raised by DHS are no longer relevant as our methodology is now consistent with the Department's recommendations in this area. This includes:

- Evaluation of the contact profiles on a double-blind basis
- Random selection of 400 "no change" contact profiles for inclusion in the double-blind evaluation
- The use of explicit, objective criteria for classifying exceptions as showing positive change or no change as a result of DUR intervention.

Items Which We Agree With in Principle But Cannot Do as a Practical Matter

The principle issue in this area is the source of the comparison profiles which we plan to use for evaluation of "positive change" in contact exception profiles. DHS prefers that these comparison profiles be drawn from the control group, or from the exceptions/non-contact population in the pilot group. We agree with DHS in principle. Drawing the comparison profiles from either of these groups is more desirable than drawing them from the contact group. Unfortunately, these profiles are not available. The data is available only for the contact group.

We believe that by carefully constructing the comparison group with "no change" contact exceptions, a reasonable assessment of the direct impact of the TDUR program can be made. We point out that it was the original intent of VCC and CCG to do a similar analysis, drawing from either the control or exception groups as a whole for comparison to the contact group. It has been our intent to use the most reliable data available to perform our analysis, and we consider our approach to be a "next best" alternative.

Elements of Our Approach Which Are Consistent With or Similar to the Original TDUR Analysis Approach

The original TDUR analysis approach was described thoroughly in a proposal from CCG to DHS on March 15, 1989. This document described at length the complex

needs of the analysis of TDUR. It effectively described why both an exceptions-level analysis and population-level analysis was necessary; the first would measure the direct impact of the TDUR program while the second would measure the indirect impact. The measure of the direct impact would be used to consider the reliability of the measure of the indirect impact.

The Department, VCC, and the new evaluation subcontractor, SRI, subsequently decided to abandon the exceptions-level measures and analysis. This was done for several reasons, including:

- Reporting and data problems at the exceptions-level
- Absence of sufficient control group data for exceptions-level comparisons
- Arguments with the exceptions-level measurement concepts.

As a result, the Department will only analyze population-level data in determining the impact of the TDUR program.

This is an area in which we agree with the original VCC/CCG proposal. Our methodology includes a cohort analysis of exceptional recipients in two groups: the contact, "positive response" group, and the contact "no response" group. Both are subgroups of the exceptions group.

We believe that unless there is a measurable, direct impact from the TDUR program, any reported indirect impact remains unverified. At this time the only way to measure direct impact is to examine individual exception review/re-review profiles for the contact group, as we are doing.

We recognize that DHS has invested a great deal of time and effort in the TDUR program and that they are concerned that any evaluation of the effectiveness of the program be reliable and thorough. We have endeavored to do just that. Our methodology was designed with consideration for quality and thoroughness within the constraints of our project's schedule and budget. If we had the benefits of the ideal evaluation budget, timeframe, data, and control-pilot group structure, certainly we would prefer to do things differently. But, we do not have this luxury and probably never will. Therefore, we are being pragmatic and focusing on what we believe is a critical question facing the Legislature: On a direct basis, what did the TDUR pilot project accomplish? We believe our methodology responds to this question with at least as much reliability as any other approach feasible within existing budgetary and scheduling restrictions.

* * * * * * * * * * *

We have endeavored to respond to the DHS letter as fully as possible so that the rationale for our approach is clear. If you have any questions regarding this material, please call me at (916) 443-6756.

Very truly yours,

ERNST & YOUNG

Perry Spiropoulos
Project Manager

APPENDIX C

THEORY AND DESCRIPTION OF TDUR

ERNST & YOUNG

APPENDIX C
THEORY AND DESCRIPTION OF TDUR

Utilization review has been a part of our country's Medicaid program since the development and issuance of the Medicaid General System Design in 1971. State Medicaid systems have been required to support Surveillance and Utilization Review Subsystems (SURS) since this document's issuance by the U.S. Department of Health and Human Services (DHHS).

SURS provide for retrospective evaluation of Medicaid services over time, giving state agencies the ability to exercise control over abuse or fraud of the program by recipients and providers of medical services. This is essentially accomplished by providing a mechanism to identify outliers with respect to normative measures of treatment and utilization, and by providing increasingly greater levels of service and utilization detail for investigating these outliers. Utilization review systems have traditionally dealt with all service types including prescription drugs.

To understand the context of California's Therapeutic Drug Utilization Review pilot program, it is necessary to understand how TDUR evolved from the traditional SURS. This section of our report provides this historical background, and proceeds to describe the California pilot program in more detail. An understanding of the historical context makes it easier for the reader to consider and assess the strengths and weaknesses of California's pilot TDUR program, which is essentially an experiment. Nevertheless, the results of the experiment and the findings related to its effectiveness are important because of recent federal legislation which mandates statewide DUR programs in all states by 1993. Knowledge gained from the California experiment will be helpful in shaping an effective statewide program in California.

This discussion is organized under the following subsection headings:

- *Concept and Application of Utilization Review Techniques*
- *The Unique Requirements of DUR*
- *Benefits Anticipated from a DUR Program*
- *Description of California's TDUR Program.*

A. CONCEPT AND APPLICATION OF UTILIZATION REVIEW TECHNIQUES

Utilization review began as a means of controlling abuse and fraud in the Medicaid program. This process, though, is not unique to Medicaid, as health maintenance organizations, provider groups, and private health insurance carriers use this process to limit overutilization. As the evolution of these concepts of fraud, abuse, and overutilization, the SUR system in Medicaid began to focus on accidental or inadvertent overutilization as well as intentional misuse of the program. As knowledge of the capabilities of SURS became more widespread, there resulted an awareness of its capabilities for evaluating long-term trends in medical utilization and its value as a tool for establishing newer and more effective policies in Medicaid coverage and reimbursement.

SURS relies on computerized systems to process and evaluate large volumes of paid claim history records. Data is generally arrayed by month of service over a nine to fifteen-month evaluative period. This large volume of data, which in California can include hundreds of millions of medical service records, must be grouped into subpopulations with like characteristics for evaluation. The highest level in this hierarchy of subpopulations includes recipient-based analysis and provider-based analysis. This means that generally the same historical data is arrayed in two ways. In the first, one can evaluate provider service and billing patterns; in the second, one can evaluate recipient utilization patterns. This is necessary because abuse of the Medicaid program can result from recipient or provider action.

At the second level of the subpopulation hierarchy, provider-based data is further categorized by type of provider (for example, hospital inpatient, outpatient, and drug) and by specialty for office-based physicians. Recipient-based data is further categorized into aid codes. These include, for example, Aged, Blind, Disabled, Aid to Families with Dependent Children, Medically Indigent, and Refugee or Alien. These categories can be further refined into aid code subsets and recipient age categories.

At the third and final level of the subpopulation hierarchy one is actually dealing with measurement criteria. Various data elements on the medical billings themselves are tested and compared to the measures which ultimately identify potential misuse of the Medicaid program. These measures may include:

- Number of unique recipients seen by a physician in a month
- Number of medical services performed in a month
- Average number of medical services performed in a day
- Average number of medical services performed per unique recipient
- Average number of prescriptions per recipient
- Numbers, costs, and ratios of types of services performed by procedure and/or diagnosis code
- Number of referrals.

These are only examples, but they serve to illustrate the nature of the data assessed in the utilization review process. In actuality, California's SURS includes thousands of different measurement criteria. Many are similar to these, and many are more complex.

The intent of the computer-generated reports is to assign a normative range of acceptability for all measures. Those providers or recipients who fall outside this normative range (outliers) are researched in more detail to determine whether in fact the historical performance was, or was not, justified.

This process is by its very nature a quantitative process. Indicators of abuse are quantitative indicators. The assumption is that excessive numbers in any of the measurements is suspect. States have been very successful in curbing abuse and fraud with these tools and techniques. A natural deficiency of this process, however, is the qualitative aspect of care.

It has long been believed by the medical community and Medicaid programs that improving quality of care will reduce medical utilization and costs in the long term. It has been the focus of DUR programs, as separate and distinct entities from SURS programs, to focus on the therapeutic (or qualitative) aspect of prescribing drugs rather than the quantitative aspect reflected in a concern over volumes of prescriptions, or abuse of certain kinds of drugs in large numbers. In contrast, California Department of Health Services (DHS) in its DUR program set out to utilize both quantitative and qualitative criterias to meet its objectives of evaluating and enhancing the therapeutic outcome of changes prescribed for persons eligible for Medi-Cal.

B. THE UNIQUE REQUIREMENTS OF DUR

Drug utilization review has been associated with Medicaid for many years. This has been the case for several reasons. First, the relatively rapid rate of release of new drugs onto the market makes it difficult to determine what may or may not be appropriate in terms of prescription quantities and/or combinations. Second, insufficient data is generally available in the case of most new drugs to determine potentially dangerous reactions resulting from drug therapy with respect to all but a few controlled circumstances as evaluated by the manufacturer. Third, the risk to the recipient from adverse drug reactions may be high as a result of the limited knowledge available about many new drugs. In an article published in Volume 308, No. 24, of the *New England Journal of Medicine*, Dr. Jerry Avorn and Dr. Stephen Soumerai state: "Less than optimal prescribing decisions by physicians are not uncommon. Drugs that are largely ineffective are used widely; dosages or combinations are often pharmacologically irrational; useful older drugs are passed over in favor of newer agents that are no more effective but may be considerably more expensive or toxic." For these reasons, the qualitative, or therapeutic, aspects of a DUR program are considered more valuable with prescription drug services than with other types of services.

Therapeutic DUR (TDUR), as in California's pilot, is differentiated from traditional utilization review by its approach to working with the provider community. To cause the kinds of therapeutic changes desired by the Medicaid program such that health is improved and the costs of illness, particularly inpatient hospital costs, are reduced overall, it is important to have the cooperation of physicians and pharmacists in the community. TDUR still relies on computer programs to identify potential drug combination or therapy problems, but because of the very care-oriented nature of the solution, this kind of a program must rely on the cooperation of the physician and pharmacist to be successful.

TDUR programs can be characterized as retrospective or prospective, and generally intervention-oriented. Retrospective DUR relies on the computer to produce historical, after-the-fact profiles of medical care for recipients whose combination of personal characteristics (age, sex, etc.) and drug utilization exceed the criteria of normality established by medical professionals. These medical profiles are screened by a group of physicians and pharmacists usually located in the community to which the serving physician and pharmacist belong. Profiles which, as a result of this screening, continue to represent a therapy anomaly result in letters or personal contact with the providers of service. In some cases it is found that the pattern of therapy is reasonable. In others, the

physician may be provided with information he or she did not previously have available. Availability of this new information could result in a change of therapy. Some differing elements may be represented in DUR programs around the country, such as whether the contact is in writing or in person. But essentially the process is the same. California's TDUR pilot program is a retrospective program.

Prospective DUR relies on providing the physician and/or the pharmacist with historical information and drug therapy information **before** the prescription is written, or before it is filled. This generally requires that the provider have access to a computer terminal in the place of work which can access recipient history data as well as drug therapy parameters. Prospective DUR will result, where appropriate, in a change of therapy before the potentially inappropriate therapy is implemented. Prospective DUR is generally applied to all prescriptions at either the physician's or the pharmacist's place of business. The process becomes more cost effective if it is combined with electronic claims submission at the pharmacy. This is because the electronic link can be used for both the UR and claim submission function, avoiding significant paper claims handling cost.

Prospective DUR should not be confused with prior authorization of drug prescriptions. Prior authorization of drugs is generally used with a pseudo closed Formulary and represents a means of seeking approval for prescribing non-formulary drugs. The process used by Medi-Cal in California, for example, requires prior authorization for non-list drugs (the list is similar to the Formulary with the exception that in California, all the drugs on the drug list are there as a result of a contract between the State and the drug manufacturers). Drug prior authorizations for non-list drugs account for less than five percent (5%) of all Medi-Cal drug services. Also, the reviewer of the prior authorization request does not consider qualitative or quantitative utilization review criteria in making his or her decision to approve or deny the authorization.

Intervention-oriented TDUR programs are simply those which rely on some sort of contact with either the physician or pharmacist of service. It is anticipated that as a result of the intervention or contact, therapy will change. Though intervention may result in direct benefits with the recipient on whose behalf contact was made, it implies a second order of benefits associated with what is referred to as the "sentinel" effect. That is, the intervention on behalf of the first recipient may result in a change of therapy to other recipients of the same physician; and again, from word of mouth, may result in changes of therapy to recipients of other physicians who were not contacted under the auspices of the program.

C. BENEFITS ANTICIPATED FROM A DUR PROGRAM

The desired benefits from any utilization review program are ultimately a reduction in overall health care costs. It is generally acknowledged that this means either making the population healthier or reducing fraud and abuse. Within the context of TDUR, improved therapies and treatments should result in a reduction in the number of people placed in health care institutions (hospitals and long-term care facilities) and a reduction in the duration of those placements. For example, program administrators at the State of Maryland, where such a program has been in operation for a number of years, relate that special evaluations performed on their Medicaid records reported a downturn in hospitalization after just six months to a year of DUR operations. The reader should be cautioned, however, that in many cases evaluative reports are completed with less than rigorous experimental techniques. Also, many analyses are done without control groups, and without proper analysis to minimize the effect of exogenous variables.

Some administrators of DUR programs are aware of these drawbacks, and though they advise against accepting results at face value, they may feel strongly that intervention-based programs have achieved results in the form of a reduction in hospitalizations. This is the case in the State of Iowa where a report was issued by the Iowa Pharmacists Association in May, 1990. This Association contracted with the State of Iowa to operate their DUR program and reported a reduction in hospitalizations. On the other hand, how reliably this reduction can be attributed to a DUR program versus other external variables is unclear because analysis has not often been rigorous enough to isolate the effects of external variables.

According to research performed by the U.S. Senate Subcommittee on Aging, there has been no known analysis where control groups have been used during the study to minimize the effect of exogenous variables. State of Wisconsin representatives related to us that they have been using a quasi-control group in their DUR program as a benchmark to evaluate the impact of the program. The control group is actually comprised of physicians and recipients in the same community where contact with other physicians is in progress. This raises the possibility that the control group physicians are changing their therapeutic patterns as a result of the sentinel effect. Even with consideration for this, Wisconsin believes its program is showing some positive results.

The need for caution in interpreting study results is further evidenced by a report prepared by the State of Florida, Office of Evaluation and Management Review, Department of Health and Rehabilitative Services, published in May of 1984. This report cited evidence of a 31 percent decline in hospitalizations among elderly clients who were at risk of therapy induced illness. This was virtually matched by a 29 percent decline in hospitalizations among those who were not at risk of drug therapy complications and therefore would not have been subject to either the direct intervention or indirect influence of the project. This demonstrates that other exogenous factors can have as much influence as the DUR project upon changes in hospitalization rates within a study group. Also, this illustrates the need for a properly managed multi-variate analysis to isolate the effects of exogenous variables and adjust resulting measures.

The California Department of Health Services, in the original implementation of its TDUR pilot program, did an effective job of establishing a control group which would have little likelihood of being affected by the pilot. Pilot and control counties were selected because of similarity in demographics and regional separation. Also, it was the Department's intention to perform a multi-variate analysis of data. Unfortunately, problems with the reporting of exceptions-based data, particularly in accuracy, made it impossible to complete the multi-variate analysis, or to perform it at an exceptions-based level. Had the multi-variate analysis been completed, as planned, it would have represented one of the more rigorous study attempts in the country, as most states with DUR programs perform a simple before and after comparison of data to estimate program results.

The most accurate analysis is one in which the DUR data is evaluated at three levels. The first is at the population level. This describes the overall movement patterns of participation and utilization within the population universe. The second is at the "exceptions" level. Groups of recipients, or subpopulations, who share the same therapeutic and demographic conditions can be analyzed for long-term changes in therapeutic patterns. The third is at the case level, where individual cases can be assessed randomly to determine if improvement can be attributed to the program. At this level, a qualitative assessment of medical data based upon expert judgement represents an alternate, or additional, means for accomplishing the analysis at an "exceptions" level which allows the analyst to determine if improvement can be attributed to the program. With each level, control groups are important. Data and findings gathered at each level can be used to correct and refine each of the other levels, and the overall study.

D. DESCRIPTION OF CALIFORNIA'S TDUR PROGRAM

California's TDUR program was implemented in pilot mode in three counties - Sacramento, Kern, and Stanislaus. DHS was responsible for administrating the pilot program. The Department's goals for the pilot were to test the hypotheses that (1) the use of a DUR system would reduce overall expenditures in the Medi-Cal program, and (2) the DUR system would reduce overall expenditures for drugs more than the existing Medi-Cal Closed Formulary/Prior Authorization system without DUR. DHS set up a control group for this experiment which included Medi-Cal eligibles in San Diego and Fresno counties. The pilot and control counties were selected based on similarities to each other and the need to prevent contact between the groups.

The purpose of this subsection of our report is to provide a detailed description of the TDUR program implemented in California, highlighting areas where the program, as implemented, deviated significantly from original program plans. These deviations are important because they limited the completeness of data produced by the program and hence our ability to evaluate its cost-effectiveness. This discussion is organized into the following topic headings:

- *Program as Originally Planned*
- *Program Actually Implemented and Deviations from Plan*
- *Implications of Deviations of Program from Planned Approach.*

1. Program as Originally Planned

In accordance with Assembly Bill 2606, DHS prepared a 40-page Request for Proposal (RFP #88-136) issued in June of 1988. The RFP described the following key contract dates and reporting requirements:

- Contractor commences implementation on January 1, 1989
- Baseline statistical data to be developed using Medi-Cal paid claims history from July 1, 1987 through June 30, 1989

- Pilot TDUR process to be fully operational by July 1, 1989 and concluding on June 30, 1991; a 24-month pilot
- Monthly administrative reporting to DHS and the Master Therapeutic Drug Utilization Review Committee (MTDURC)
- Quarterly statistical reporting to DHS, the MTDURC, and the Legislature
- Final pilot evaluation report due to DHS and the State Legislature by July 15, 1990.

As stated in Section I, the contractor selected to implement the TDUR pilot and carry out the requirements of the RFP was The Virginia Computer Company (VCC). The basic components of the TDUR pilot, and the major responsibilities of VCC in implementing the program, were:

- *Conduct Pilot Program Operations*
- *Operate TDUR Computer Processing System*
- *Conduct Pilot Program Evaluation Activities.*

Our discussion of program plans will focus on these three areas.

Pilot Program Operations

VCC's pilot program operations responsibilities were described in the DHS RFP. These duties were to include:

- Use of an orderly project protocol for the DUR process of computerized inquiry and discovery of cases of drug use which require attention
- Manual evaluation of those cases by medical professionals
- Follow-up provider notification
- Referrals of cases to the Department and/or licensing boards when warranted.

Standards for carrying out these duties were to be set by the Master TDUR Committee. Tasks within the DUR process which were to be guided by the MTDURC included:

- Establishment of a computer system capable of tracking drug utilization by drug, illness, beneficiary and provider
- Development and implementation of policies for controlling inappropriate, unnecessary, unjustified or expensive drug utilization
- Standards used to assess cost, dosage, duration, and effectiveness of medications relative to a particular diagnosis
- Definition of inappropriate drug utilization and/or costs and protocol for informing providers thereof.

In addition, VCC, under the guidance of the master committee, was to develop the quantitative and qualitative exceptions criteria to be used for the DUR process. The quantitative criteria were to be designed to identify and act on *select* prescribing and dispensing characteristics which result in unnecessary expenditures for drug products. The qualitative criteria were to assist in identifying and acting on *select* prescribing characteristics which cause therapeutic complications requiring costly remedial care.

TDUR Computer Processing System

The functional and technical requirements for the TDUR computer processing system were not defined in the RFP. DHS specified only that VCC would design, develop, install and operate "a computer-based DUR system", and that the system must be compatible with Medi-Cal's data processing systems so that "all data elements" required for the DUR process could be extracted from the Medicaid Management Information System (MMIS) program history files.

Although data requirements were not specifically defined, reporting requirements were described in detail. Therefore, required data elements could be inferred from the report descriptions. Report formats were to be cumulative to-date by month of service, and were to show by aid category the comparisons on current data with MMIS baseline data. Quantitative and qualitative comparisons were to be made for both the pilot and control groups. Qualitative comparisons were to be reported for:

- Incompatibilities of medications to diagnosis

- ❑ Overprescribing and underprescribing of drugs
- ❑ Use of contraindicated and incompatible drugs
- ❑ Drug-induced physician visits, outpatient hospital visits, emergency room visits, clinic visits, hospital admissions, skilled nursing facility (SNF) admissions, and intermediate care facility (ICF) admissions
- ❑ Days of drug-induced hospitalizations and institutionalizations
- ❑ Mortality
- ❑ Morbidity
- ❑ Overall impact of qualitative DUR on program utilization of, and program costs for, each of the services specified above, as measured against MMIS baseline and against the control group.

The quantitative DUR was to be measured by comparisons of, and trends in:

- ❑ Overutilization of prescribed drug(s)
- ❑ Pharmacy overutilization of prescribed drugs with reference to dispensing practices
- ❑ Physicians overutilization of prescribed drugs with reference to prescribing practices
- ❑ Overall impact of quantitative DUR on the program utilization of, and program costs for, drugs, as measured against MMIS baseline and control.

Pilot Program Evaluation Activities

In its proposal to the State, VCC described its plan to subcontract pilot program evaluation activities to Consolidated Consulting Group (CCG). Subsequent to the VCC award announcement, significant contract negotiations occurred. These negotiations resulted in contract amendments, addressing requirements for State ownership of the systems software and therapeutic standards after conclusion of the project and for VCC to clarify its work plan relating to its responsibilities for pilot "program evaluation."

With regard to the latter, VCC submitted a detailed technical proposal to DHS on March 15, 1989. This workplan clarified the contractor's approach to evaluation of the TDUR pilot project. Our description includes a summary of the following elements of the planned approach:

- *Objective of the Pilot Evaluation Plan*
- *Design of the Pilot Evaluation Plan*
- *Measuring the Direct and Indirect Effects of TDUR*
- *Quarterly Summary Analyses.*

- *Objective of the Pilot Evaluation Plan*

  The objective of the planned evaluation was to "measure the effect of the TDUR program on health care expenditures and utilization by Medi-Cal non-contract eligibles." Non-contract eligibles are those eligibles who do not participate in prepaid health plans. This objective was to be achieved by "producing quarterly analyses of the information contained in the (operations) reports generated by VCC and conducting a multi-variate regression analysis after the program has been in operation a sufficient time to generate the data required for analysis." These quarterly analyses would evaluate the "direction and general magnitude of changes in health expenditures and utilization" in the TDUR and control populations. The purpose of the regression analysis was to "specify the degree to which changes in health expenditures and utilization are attributable to the TDUR program." Clearly, the objective of the TDUR program evaluation as planned included utilizing a methodology which would justify attributing any benefit measured in the TDUR population to the TDUR program.

- *Design of the Pilot Evaluation Plan*

  The pilot evaluation design was characterized as "an interrupted time-series, quasi-experimental design." Quasi-experimental meant that California's Medi-Cal eligibles were not randomly assigned to the TDUR and control populations as they would be in a true experiment. Rather, existing groups of individuals already living in the pilot and control counties were "assigned" to the TDUR and control populations, respectively.

  The thrust of the evaluation design was to select demographically and geographically comparable pilot and control counties, to identify the factors other than TDUR

which would account for effects on health expenditures and utilization, and to measure the direct and indirect effects of TDUR in the pilot counties after controlling for the other factors. Factors which would affect health expenditures other than TDUR are called control, or independent, variables. Factors which account for the direct and indirect effects of TDUR are called outcome, or dependent, variables. Multi-variate regression analysis was the statistical tool which was to be used to isolate the effects of changes in the dependent variables from changes in the independent variables.

The independent variables included factors such as socio-demographic characteristics, medical marketplace characteristics and time trend variables.

❑ *Measuring the Direct and Indirect Effects of TDUR*

The DHS RFP required the TDUR contractor to evaluate both quantitative and qualitative measures of the TDUR program. Meeting this requirement was a crucial component in the planned evaluation approach. The pilot evaluation plan specified "quantitative" to mean measuring the direct effect and "qualitative" to mean measuring the indirect effect of TDUR. These measures were defined as follows:

- Direct Effect - Measures of changes in outcome variables for those with exceptions who received an intervention (e.g. a letter) compared to those with exceptions who received no intervention

- Indirect Effect - Measures of changes in outcome variables for the entire TDUR and control populations over time.

The direct effect, in other words, was to be determined by comparing the outcome variables for the contact group (those whose providers received letters) and a comparable group of eligibles in the control population (those exceptions whose providers would have received letters if they had been in the pilot population). The indirect effect was to be determined by comparing the outcome variables for the entire TDUR population and the entire control population. A list of the outcome variables for measuring the direct and indirect effects, as planned in the original evaluation design, is provided in Appendix F. They included measures for utilization and cost of health services including prescriptions, physician visits, hospital care, etc.

❑ *Quarterly Summary Analyses*

The Quarterly Summary Analyses to the MTDURC and the Legislature were to begin July 1, 1989. The first report was

to contain results of measuring dependent variables for the baseline period (two years prior to implementing the TDUR program.) Subsequent reports would compare trends for the TDUR and control populations, including measures for direct and indirect effects.

The plan included provisions for handling time lags in processing Medi-Cal claims data, and instability in the TDUR and control populations (i.e., people join and leave the Medi-Cal program at varying intervals). Also considered were methods for generalizing the evaluation results to all Medi-Cal eligibles. A number of reports were planned by CCG. These reports were to be organized as follows:

- Quarterly Utilization Analysis which was to examine the effects of TDUR on utilization of health care services by those Medi-Cal non-contract eligibles with identified exceptions and the entire population of Medi-Cal non-contract eligibles.

- Quarterly Expenditure Analysis which was to examine the effects of TDUR on State expenditures for health care services for both those Medi-Cal non-contract eligibles with identified exceptions, and for the entire population of Medi-Cal non-contract eligibles.

- Provider Exception Analysis which was to examine the effects of TDUR on provider behavior by tracking changes in the percentage of recipients with exceptions per provider over time.

- Therapeutic Intervention Analysis which was to track changes in the number of exceptions per 1,000 Medi-Cal eligibles after intervention by the TDUR program.

Had these quarterly reports been produced as planned, they could have formed the basis for the legislatively mandated independent evaluation (i.e., this evaluation study by Ernst & Young).

2. Program Actually Implemented and Deviations from Plan

When the time comes to actually implement a large scale research project such as this, there are always a number of unplanned events which require deviations from the original project plans. Such is to be expected, and such was the case with the TDUR program implementation. However, a number of deviations which occurred during the

pilot program implementation have a direct bearing on the ability to adequately assess its cost-effectiveness. In this section we first describe the actual implementation of TDUR under this program. This description provides an overview of the process used from the inception of TDUR to-date. It serves as a frame of reference for consideration of the deviations from plan. The discussion of deviations follows the description of the actual implementation.

The TDUR program in California was implemented in three pilot counties - Sacramento, Kern, and Stanislaus. Based on similarities to the pilot counties, San Diego and Fresno Counties were selected as control counties, in order to provide a basis for comparison of therapeutic outcomes. The total combined number of Medi-Cal eligibles in the pilot counties is roughly 300,000, and constitutes the TDUR population. The roughly 300,000 Medi-Cal eligibles in the control counties make up the control population.

According to DHS, thirty percent (or 90,000) Medi-Cal eligibles within the TDUR population receive prescription services and may therefore experience problems with drug therapy. These 90,000 eligibles represent the benefit group, or group of Medi-Cal eligibles who could potentially be affected by any direct or indirect benefit of California's TDUR program.

During our evaluation time frame, the fourteen month time period from July 1, 1989 through August 30, 1990, the TDUR computer system reviewed Medi-Cal paid claims for both the TDUR population and the control population. Exceptions (eligibles whose paid claims showed a potential problem with drug therapy based on the master file exception criteria) were identified, and statistical exception data was accumulated for both populations. However, **exception profiles** (reports listing all paid claims for an exception for the previous 90 days) **were only produced for the TDUR population,** not for the control population.

As reported by VCC, almost 17,000 exception profiles were generated during our fourteen month evaluation time frame. All of these profiles were reviewed by local DURC members. VCC was responsible for recruiting physicians and pharmacists from each county to participate in the committees. These medical professionals were given an honorarium of $100 per month for their services. In total, members were requested to review an average 1,200 recipient profiles per month.

During the initial review, the DURC members annotated the exception profiles with actions required, as follows:

- ❑ Close case with no further action
- ❑ Send a letter to the provider
- ❑ Request case history for more information
- ❑ Refer to Ad Hoc Committee.

Case referrals to the Ad Hoc Committee could result in case closure or referral back to the DURC with instructions for further action. Under unusual circumstances, the Ad Hoc Committee would require case referral to the Department of Health Services when fraud and abuse was suspected, to the provider's licensing agency or professional society for peer review, or to the state MTDURC.

Annotated exception profiles were used as input documents to the automated system. Exception cases were tracked in the Patient Tracking Module and remained open, or in suspense, until resolved. Under unusual cases the DURC members or the VCC pharmacy manager would contact providers to discuss the case problem.

The DURC members found that many of the exceptions did not represent true drug therapy problems, but rather problems with the master file exception criteria. Therapeutic criteria specifications were defined by VCC, approved by the MTDURC, and then held constant for the 24 months. **Exhibit C-1**, on the following page, presents the eight fixed rotating therapeutic criteria categories used or planned for use during the complete pilot. These groupings were arranged to facilitate the county DURC review process. On a rotational basis, one of these eight therapeutic criteria categories was selected as the focus of each monthly exception process. Pertinent articles from medical journals and other sources covering the month's topic were prepared by the VCC Pharmacy Manager, who managed the meetings.

TDUR Exception Processing Criteria and Schedule

CRITERIA	THERAPEUTIC CLASSES INCLUDED	MONTHS USED
1. Anti-Ulcer Preparations	Class 01	7/89, 3/90, 11/90
2. Anti-Arthritic	Class 40, 41, 42	8/89, 4/90, 12/90
3. Drug to Diagnosis Incompatibilities	All Classes (There are a total of 99 Therapeutic Classes)	9/89, 5/90, 1/91
4. Duiretics & Hypertensives	Class 53, 70, 71, 72, 73, 74, 76, 79	10/89, 6/90, 2/91
5. Anti Depressant	Class 11	11/89, 7/90, 3/91
6. Tranquilizer/Anti-psychotic, Hypnotic, etc.	Class 07, 08,09, 20, 46,47,48, 49, 58	12/89, 8/90, 4/91
7. Oral Contraceptives	Class 36, 63	1/90, 9/90, 5/91
8. Overutilization	All Classes	2/90, 10/90, 6/91

About 13 percent of the total exceptions produced during the pilot (2,200 exceptions) represented what the DURC believed to be true drug therapy problems, and the providers of these eligibles were contacted by letter to advise them of the problems. **Exhibit C-2,** following Exhibit C-1, shows a sample provider letter of the type which was used in the TDUR pilot. The 2,200 eligibles whose providers of service were contacted represent the group of Medi-Cal eligibles for whom we could potentially measure any direct benefit from California's TDUR program. This group is called the **Contact Group** in this report.

Four to five months after providers were contacted, a Re-Review Profile (second report listing paid claims for the previous 90 days) was automatically generated for each exception. Each re-review profile was compared to its original to determine if the identified problem had been corrected. Exception cases could then be closed, the provider recontacted, or problem cases could be referred to the Department of Health Services for special action if this appeared necessary.

The DURC members determined if the re-review case should be closed or if further action was warranted. The action required was also annotated on the re-review profile. If the re-review case was closed, the VCC Pharmacy Manager indicated the type of closure, annotating on the re-review profile to indicate the observed behavior pattern, as follows:

- ❑ Provider (prescriber/dispenser) therapy change
- ❑ Patient utilization change, e.g., under care of a single or limited number of providers
- ❑ No change that appears to be related to TDUR activities.

VCC reported that the re-review profiles of about 400 exceptions showed that they had improved in the identified problem area. The exceptions who showed improvement in the identified problem after their providers were contacted are called the **Response Group** in this report, and the exceptions who did not show improvement in the identified problem after their providers were contacted are called the **Non-Response Group.**

SAMPLE INITIAL PROVIDER NOTIFICATION LETTER

MEDI-CAL THERAPEUTIC DRUG UTILIZATION REVIEW COMMITTEE FOR THE STATE OF CALIFORNIA

(Date)

(Provider Name) (Provider ID); (Drug ID)
(Provider Address)

RE: (Patient Name), Medicaid ID: (Patient ID)

Dear Pharmacist:

The California Department of Health Services has implemented a Therapeutic Drug Utilization Review (TDUR) Program designed to monitor the use of prescription drugs by Medi-Cal recipients. This new program is being initially implemented in Kern, Stanislaus and Sacramento counties. The goal of the TDUR program is to provide practitioners with clinically relevant data to help assist them in managing patient outcomes.

The TDUR Program routinely profiles patterns of prescription drug use in the Medi-Cal patient population in the above mentioned counties. Applying therapeutically oriented criteria recommended by the Medi-Cal Therapeutic Drug Utilization and Review Committee, this program attempts to identify instances of drug utilization which may require additional review. These drug histories then undergo a confidential review by a County Committee composed of local physicians and pharmacists.

The enclosed profile of your patient indicates that the patient is receiving several prescriptions prescribed by different physicians. Since we are unaware if all providers involved are aware of the drugs prescribed by other physicians, we are providing this profile to all physicians and pharmacists involved for evaluation and action, if appropriate.

It would be most helpful to the County Committee if we could learn from you if this information has been useful, as well as any specific comments or changes to therapy you may take concerning this particular patient.

Please feel free to use a copy of this letter to simply note your comments and return in the enclosed envelope. You may keep the enclosed profile for your records. Thank you for taking the time to respond.

Sincerely,

(Signature)

Medi-Cal TDUR Manager
for the Medi-Cal TDUR Committee
Administered by The Virginia Computer Company

Table C-1, below, summarizes the TDUR activities over time, in terms of numbers of profiles, contacts, responses, etc., as reported by VCC in its monthly activity reports.

Table C-1
TDUR Activity Summary for Pilot Counties

	FY 1990 7/89-6/90	FY 1991 7/90-8/90*	% of Total	Total
Exception Profiles	14,636	2,111		16,747
Profiles With Letters Requested by DURCs	1,998	211	13% of line 1	2,209
Profiles With Change	396	41	20% of line 2	437
Profiles With Responses To Letters	906	134		1,040
Letters Requested by DURCs**	7,210	719		7,929
Letters Delivered**	5,506	493		5,999
Responses Received**	1,578	233		1,811

* Due to time lags, data available for FY 1991 is limited to two months. These data were combined with data from fiscal year 1990 to constitute our 14 month evaluative period.

** Typically multiple providers were contacted for each recipient. Therefore, the number of letters generated and delivered exceeds the number of profiles with letters generated, and the number of responses received exceeds the number of profiles with responses.

Our discussion of program deviations is organized as follows:

- *Deviations in Pilot Program Operations*
- *Deviations in TDUR Computer Processing System*
- *Deviations in Program Evaluation Activities.*

Deviations in Pilot Program Operations

Significant deviations which occurred in program operations occurred in criteria development and maintenance, and in provider notification. This discussion includes a description of each deviation type.

- *Deviations in Criteria Development and Application*

  Avoidance of institutionalization and cost savings, which were objectives of the TDUR program, were not specific objectives of the therapeutic criteria specifications.

  Development of the therapeutic criteria did not emphasize support of the legislated objectives of the TDUR pilot. There was a lack of MTDURC discussion of these objectives as members approved the criteria to be used, according to comments from DURC members and MTDURC minutes. The therapeutic criteria consisted of an unmodified product previously developed by VCC for use in earlier applications for other states. While these criteria may have provided a means of identifying a number of potential drug therapy problems, there was no attempt made to modify them to focus on specific medical problems or specific Medi-Cal populations which experience a higher level of institutionalization. Nor was there any focus on particular drugs or drug therapy problem types which typically result in a high percentage of Medi-Cal costs.

  Problems with criteria were confirmed during interviews with DURC members. One example cited was that anti-coagulants were not included in the criterias. Anti-coagulants are known to cause a high rate of hospitalizations if not used correctly.

  A "maximum services" criteria category was added to the eight therapeutic criteria categories which were shown in Exhibit C-1.

  The maximum services criteria was added during the TDUR pilot to produce exceptions for eligibles whose paid claims data revealed utilization of more than a pre-defined number of services during a 90-day exception period. This type of review was already performed by the state's federally-approved Medicaid Management Information System (MMIS), Surveillance and Utilization Subsystem (SURS).

  The MTDURC meeting minutes of February 27, 1989, show that the committee wanted to avoid using criteria that would duplicate SURS. This type of criteria was not required by AB 2606, and the MTDURC questioned its

appropriateness for the pilot. DHS called for the quantitative criteria in their RFP and required its use.

The county DURC members expressed a certain frustration when dealing with the maximum services exceptions. In most cases, the recipient was being abusive of the Medi-Cal program, e.g., doctor-shopping for drugs. Since the recipient's behavior in these cases was usually beyond the control of the physicians and pharmacists involved, provider notification would not appear to be an effective means of effecting change. Furthermore, these types of problems fall outside the intent of Therapeutic Drug Utilization Review, and are the responsibility of DHS fraud and abuse intervention units.

- *Deviations in Provider Notification*

  Problems with the State's Provider Master File resulted in an inability to deliver approximately 24 percent of the provider notifications requested by the DURCs

  The MMIS claims history file, used by the TDUR system to generate exceptions, contained identification numbers for providers of services associated with each recipient. In order to notify providers, their names and addresses had to be extracted from the State's Provider Master File (PMF). However, the PMF did not contain up-to-date information. Many of the providers shown on the MMIS claims history file were not on the PMF. Also, the PMF showed a number of invalid addresses. This resulted in an inability to identify providers for 14 percent of the provider notifications requested by the DURCs and an inability to deliver another 10 percent of the provider notifications, as they were returned by the post office due to invalid addresses.

Deviations in the TDUR Computer Processing System

Because of the problems observed by DHS in the exceptions-based data, it was recommended by SRI, and agreed to by DHS, to deviate from the original study design requiring the program evaluation to include exceptions-based measures. SRI will complete their analysis using only population-level data.

At the beginning of our study we engaged in several conversations with DHS staff involved in the TDUR program regarding the data reports produced by the VCC system. Essentially, it was the opinion of DHS staff that a significant amount of the data produced by the VCC system was inaccurate. The DHS Statistics Branch had attempted to validate a variety of the reports produced by VCC using other existing data bases belonging to the Department, and was unable to validate significant amounts of data. They were, however, able to confirm with some level of comfort population-level statistics produced by VCC. They were unable to validate the exceptions-based measures important to completing the analysis described in the original study design, such as the number of physician visits per 1,000 eligibles for the DUR and control eligibles with exceptions.

The Department first became concerned regarding the accuracy of TDUR data files being maintained by VCC at the MTDURC meeting of September 1990. VCC reported that computer errors were identified that resulted in an erroneous count of program eligibles. The data which were corrupted invalidated the raw data reports that were to be used in preparing the Quarterly Summary Reports for the Legislature.

Although VCC's October 1990 TDUR Administrative Report to DHS indicated that the problem had been corrected and new data produced, the November 1990 Quarterly Report, the first to be produced, continued to report outstanding problems. As a result, DHS determined that exceptions-based data for both the pilot and control groups were unreliable. These data problems are being evaluated independently by DHS, and DHS plans to compare its reports with those produced by the TDUR computer system in order to identify the inconsistencies in data. At the time of this report, however, the evaluation of this data by DHS is not finished.

Deviations in Pilot Program Evaluation Activities

VCC originally planned to subcontract its program evaluation responsibility to its affiliate, Consolidated Consulting Group. CCG produced a detailed technical proposal containing the original pilot evaluation design approved by the MTDURC. Key elements of the plan included multi-variate regression analysis to measure the

direct, and indirect, effects of TDUR while controlling for extraneous factors, and the submission of Quarterly Reports to summarize interim findings. The Quarterly Reports were to begin July 1989 by summarizing baseline data. CCG subsequently went out of business and was later replaced by Stanford Research Institute (SRI). This change, in combination with the Department's concern over data problems in the VCC system, resulted in deviations from the original pilot evaluation plan, including the following:

- There was a significant delay in selecting a replacement for CCG
- The planned evaluation design was significantly altered
- As of the publication of this report, only two Quarterly Reports have been submitted to the State Legislature, and neither contained meaningful statistical analysis of the TDUR pilot program's potential for cost savings.

Each of these deviations are further discussed below.

- *Delay In Replacing CCG*

  There was a sixteen-month delay in replacing CCG.

  In May, 1989, VCC announced that CCG would no longer be able to serve as evaluation subcontractor due to a loss of key personnel and other business issues. VCC expressed the desire to replace the subcontractor as soon as possible in order to prevent delays in implementing of the reporting elements of operations.

  Though potential candidates for the research subcontract were discussed at a June 23, 1989, MTDURC meeting, it was not until September, 1990, that SRI was selected and available to meet with MTDURC members. This represented a sixteen-month delay in replacing CCG.

- *Alteration of Original Pilot Evaluation Design*

  It has been decided to measure the cost-effectiveness of TDUR by measuring only indirect effects.

  SRI provided an overview of the suggested new evaluation design at the September 1990 MTDURC meeting. In its proposal, SRI planned to retain the outcome variables developed in the original plan to measure the direct and

indirect effects of the TDUR program. However, in an October 1990 memorandum to DHS and VCC, SRI indicated that they had identified some issues regarding the use of exceptions criteria (i.e., measurement of the direct effects of the TDUR program). This memorandum stated that measuring the indirect effects of TDUR (measures with respect to the entire population of eligibles) should be the primary focus of the evaluation, and that measuring the direct effects should be relegated to a supportive explanatory role.

Three months later, subsequent to meetings between Ernst & Young, DHS, and SRI, SRI issued another memorandum which stated that "because of technical and definitional limitations", direct (exceptions based) measures should not be included in the evaluation design. SRI strongly recommended that only indirect (population based) measures of TDUR effectiveness be used. Thus, a critical element of the original evaluation design was altered. As a result, direct effects of TDUR have not been, and will not be, measured or analyzed by SRI. SRI's memorandum described difficulties in collecting and analyzing exceptions-based data in the TDUR program that they believed would negate the potential benefit which might be achieved by measuring direct effects. SRI recommended determining cost-effectiveness of TDUR by measuring only indirect effects.

- *Failure to Produce Quarterly Reports*

There was a failure to produce all the Quarterly Reports required by the Legislature. The two reports which were produced included no analyses.

The original TDUR plan included quarterly reports describing program performance to the Legislature. These reports were a requirement of AB 2606. The reports were to include analysis of the cost-effectiveness of the program through a number of measurement variables of therapeutic drug treatment. As of the date of issuance of this study, only two Quarterly Reports have been submitted to the Legislature and these did not include the required analysis. The reasons for this deviation include:

- A delay in replacing CCG (described above)
- An apparent failure to incorporate quarterly reporting data requirements into the design of the TDUR computer processing system
- Data accuracy problems with the TDUR computer processing system.

3. Implications of Deviations From Planned Approach

Six major deviations from the planned approach were identified. These deviations are summarized below together with a statement regarding their implications on the administration and evaluation of TDUR.

Deviation #1: ***The maximum services exception criteria category was added to the program design and changed frequently during the pilot program. This category was not a requirement of the Legislation. In contrast, therapeutic exception criteria was held constant.***

Implications: The maximum services exception criteria category was not a requirement of AB 2606. It was added as a program requirement by DHS. Accommodation of these new criteria necessitated additional computer programming by VCC. Inclusion of the criteria, followed by a steady relaxing of the exception threshold, had the following effects:

- ❑ Increased the exceptions volume
- ❑ Increased maximum services exceptions as a percent of the overall volume of exceptions
- ❑ Attracted DURC members attention away from drug therapy issues to focus on program abuse issues
- ❑ Decreased the number of exception letters generated as a percent of total exceptions
- ❑ Altered the experimental nature of the original design.

These impacts are illustrated in **Exhibit C-3** through **C-6**. These exhibits include data from the period July, 1989 through December, 1990. While this includes information beyond the fourteen month evaluation period of our study, we feel the additional data provides a better picture of overall trends in the pilot program. These exhibits include:

- ❑ **Exhibit C-3**: Modification of Maximum Services Criteria During the Project
- ❑ **Exhibit C-4**: Maximum Services Exceptions as a Percent of Total ExceptionS
- ❑ **Exhibit C-5**: Average Number of Exception Profiles Generated Over Time

Modification of Maximum Services Criteria During the Project

MAXIMUM SERVICES CRITERIA

EXCEPTION PARAMETER	# OF PRESCRIPTIONS	# OF PRESCRIBERS	# OF PHARMACIES
MONTHLY SETTING:			
July 1989	60	10	7
August 1989	60	10	7
September 1989	60	10	7
October 1989	60	10	7
November 1989	60	10	7
December 1989	30	9	5
January 1990	30	9	5
February 1990	30	9	5
March 1990	30	9	5
April 1990	30	9	5
May 1990	25	5	5
June 1990	25	5	5
July 1990	25	7	5
August 1990	25	7	5
September 1990	25	6	5
October 1990	20	5	5
November 1990	20	5	5
December 1990	20	5	5
January 1991 February 1991 March 1991 April 1991 May 1991 June 1991	Not available	Not available	Not available

NOTE: A total drug cost exception of $10,000 was set, but no exceptions were ever generated.

Maximum Services Exceptions as a Percent of Total Exceptions

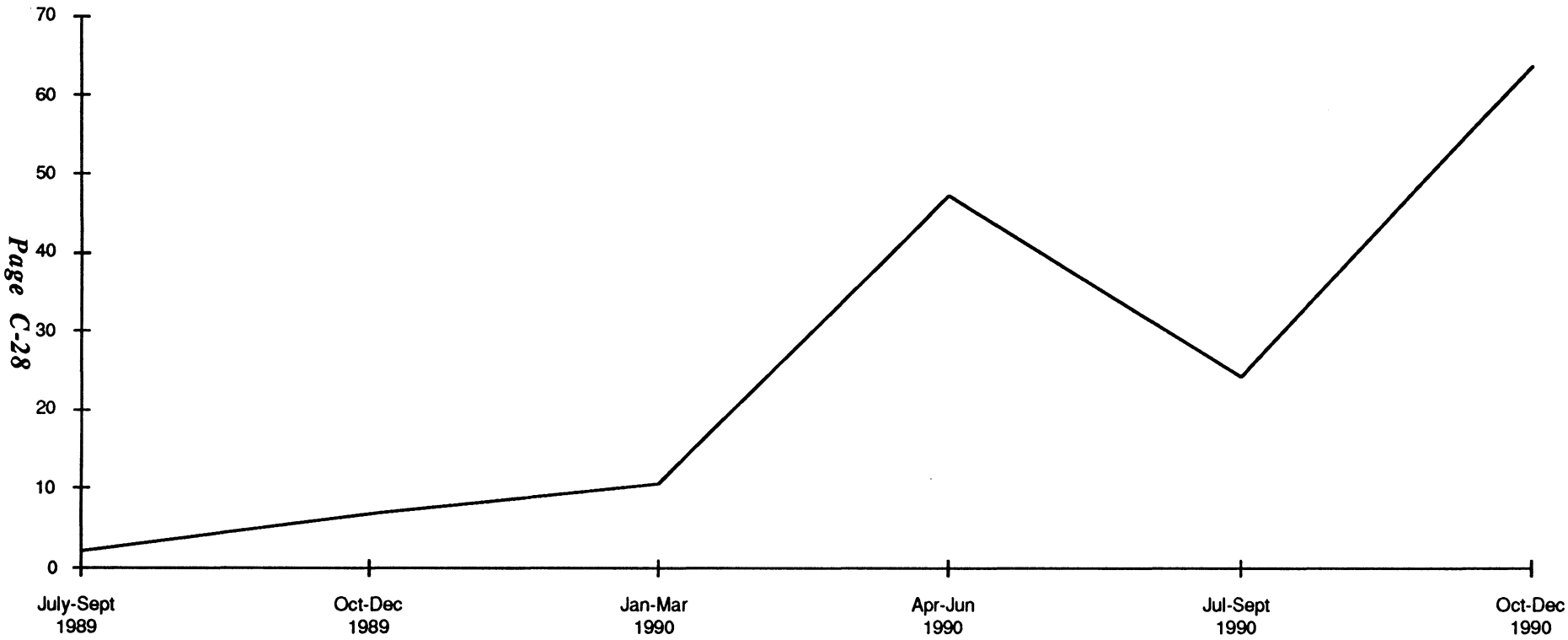
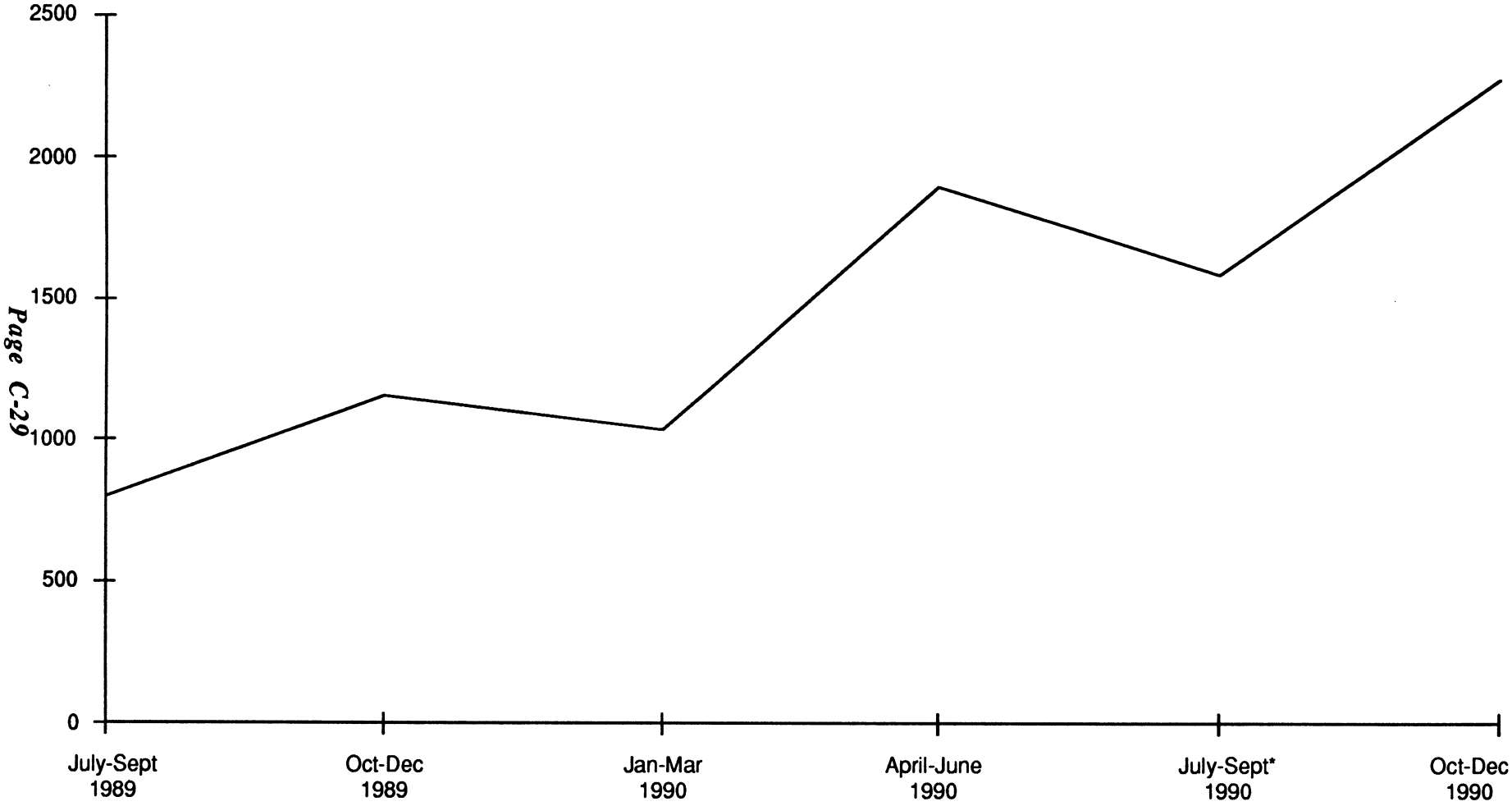


EXHIBIT C-4

Average Number of Exception Profiles Generated Over Time



* *Due to payment delays during July 1990, only slightly more than one month of history was available, which caused a temporary decrease in the number of exceptions and profiles.*

EXHIBIT C-5

- **Exhibit C-6**: Percent of Exception Profiles with Letters Over Time.

This deviation had two results. First, the decision to vary the maximum services exception level contributed to SRI's decision to abandon exceptions-based measures as a means of determining program impact. SRI determined that the variations created problems with the experimental design which could not be resolved.

Second, distraction of DURC members from therapeutic issues to issues of abuse changed the scope of the study design and may have impacted the ability of the DURCs to benefit the program as planned, particularly as this effort was redundant with existing fraud and abuse programs at DHS.

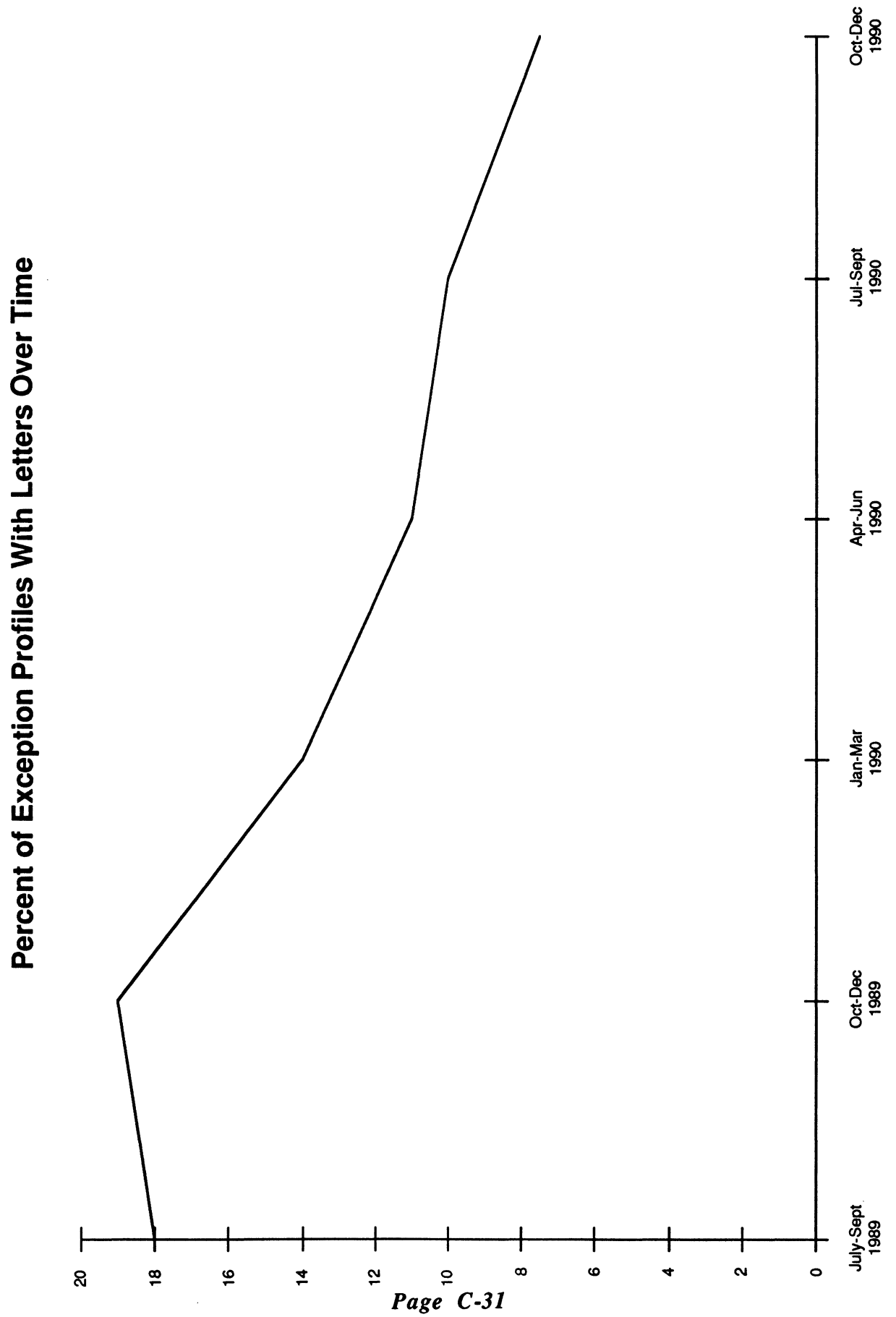
Deviation #2: ***There was a failure to consider avoidance of institutionalization or cost savings objectives when developing TDUR exceptions criteria.***

Implications: AB 2606 clearly set out to test the cost-effectiveness of a TDUR program, with special emphasis on savings achievable through improved quality of care seen through reductions in institutionalization. It appears that sufficient attention was not given to the unique criteria which would benefit TDUR in California, including:

- Implementation of a "canned" set of criteria without evaluation or change
- Lack of consideration of specialized criteria by the MTDURC
- Lack of prioritization in applying criteria which would target either drug therapies which could result in high levels of institutionalization, or the avoidance of drug therapies with little or no potential for cost saving benefits to the program.

This development may have resulted in reduced program impact and a greater focus on activities not specifically required in the legislation.

EXHIBIT C-6


Percent of Exception Profiles With Letters Over Time
20
18
16
14
12
10
8
6
4
2
0
July-Sept 1989
Oct-Dec 1989
Jan-Mar 1990
Apr-Jun 1990
Jul-Sept 1990
Oct-Dec 1990

Deviation #3: ***Problems with the state's Provider Master File resulted in an inability to deliver approximately 24 percent of the provider notifications requested by DURCs.***

Implications: The recipients in the contact group were seeing an average of four providers each during the evaluative period. Almost 2,000 provider notifications requested by DURCs for this group which should have been mailed were not. Providers who never received notification obviously could not respond to direct TDUR intervention. This by definition limited the direct effects of TDUR.

Deviation #4: ***Data accuracy problems with the TDUR computer system contributed to a decision by DHS and SRI to substantially change the methodology and scientific approach planned for evaluating the effectiveness of the TDUR program.***

Implications: Data accuracy problems with VCC reports have contributed to the lack of overall performance data for TDUR which have in turn resulted in the inability of VCC to provide quarterly reports to the Legislature regarding program benefits. This has resulted in a substantial deviation from the original evaluation plan by SRI. The absence of reliable data about the program's performance caused us to implement independent research to try to confirm the direct benefits of TDUR.

Deviation #5: ***There was a delay of 16 months in locating a replacement evaluation subcontractor when CCG withdrew from the project.***

Implications: This delay contributed substantially to the failure of VCC to meet the legislative requirement to produce Quarterly Reports. Only two reports have been produced, and these do not contain the analysis required to determine TDUR cost-effectiveness. The absence of these reports has caused us to implement independent research to attempt to confirm the direct benefits of TDUR.

Deviation #6: ***The evaluation design intended for this program has been significantly altered.***

Implications: The original pilot evaluation design thoroughly addressed the complex needs of the analysis of TDUR. It effectively described why both an exceptions-level analysis, and population-level analysis was necessary. The first would measure the direct impact of the TDUR program, while the second would measure the indirect impact. The measure of the direct impact would be used to consider reliability of the measure of the indirect impact. The implication of abandoning this approach, and measuring only the indirect impact, is that any reported indirect impact remains unverified. Further implications are that the direct impact must be ultimately measured and we believe the only way to do so, **given the limited data**, is to examine individual exception review/re-review profiles for the contact group.

APPENDIX D

EVALUATION METHODOLOGY

ERNST & YOUNG

APPENDIX D
EVALUATION METHODOLOGY

The overall goal of this evaluation, as set forth in AB 2606, was to determine the cost-effectiveness of the TDUR program and to develop recommendations regarding its continuation. Specific evaluation objectives related to this goal were as follows:

- *To evaluate the impact on institutionalization of Medi-Cal eligibles by operation of the Medi-Cal therapeutic drug utilization review process*
- *To evaluate the cost impact of the Medi-Cal therapeutic drug utilization process*
- *To determine whether the TDUR system is cost-effective*
- *To determine whether the TDUR system should be continued.*

This section describes the methodology we used to complete this study. The discussion is organized into the following topics:

- *Constraints on the Evaluation Framework*
- *Description of the Methodology*
- *Summary of Rationale for the Selected Approach*
- *Conclusions.*

A. CONSTRAINTS ON THE EVALUATION FRAMEWORK

The most significant constraint on our evaluation framework was the limited availability of information to address the effects of TDUR. It was a requirement of AB 2606 and the contract between DHS and VCC that VCC would produce evaluative data on a quarterly basis as an output of the TDUR computer processing system. It was further stipulated that the data would be analyzed to determine the effects of TDUR, and the analyses would be summarized in Quarterly Reports to the Legislature. Had this been done, the scope and objectives of this independent evaluation would have been limited to reviewing reports summarizing the effects (expected benefits) of the pilot TDUR program,

and extrapolating from those benefits the projected cost-effectiveness of such a program in California.

The two Quarterly Reports submitted to-date do not assess TDUR benefits for the pilot program, which was limited to three counties, and cannot be used for extrapolating cost-effectiveness of such a program on a statewide basis. Furthermore, the data required to measure TDUR benefits, as originally planned, has not been produced. Producing these data is clearly outside the scope of this evaluation. Therefore, it was essential that we devise a feasible alternative to provide an adequate measure of direct benefits attributed to TDUR. This discussion describes the alternatives considered for measuring the direct benefits attributed to TDUR and completing our study within the directive of AB 2606.

Given that the data does not exist to support the evaluation as originally planned for this program, we defined the available alternatives for meeting our objectives by the data available under this program. The original evaluation design called for measuring both direct and indirect effects of TDUR. The research being conducted by SRI has the potential to measure indirect effects, but does not address direct effects. In our opinion completing an analysis of TDUR benefits required measuring direct program benefits.

Direct program benefits could only be measured by examining the treatment or Contact Group within TDUR, since this is the only group which experienced the intervention. Our problem in this endeavor was identifying an alternative comparison group since exception profiles were not generated for the comparable Control population. Alternatives for a comparison group were limited to exceptions *within the TDUR population* since no data existed for exceptions in the Control population. The available alternatives included:

- *Response Group alone*
- *Response Group versus Exceptions not contacted*
- *Response Group versus Non-Response Group.*

Before further describing our methodology and rationale, we provide the following definitions of terms used in our study report:

- *Eligibles* -- refers to all Medi-Cal recipients in both the pilot and control counties (about 300,000 in each)

- *Exposed Eligibles* -- refers to Medi-Cal recipients in both the pilot and control counties who used drugs (about 90,000 in each)

- *Exceptions* -- refers to the number of exceptions generated by the computer based upon specific drug utilization criteria (about 17,000 in the pilot counties; not applied to the control counties)

- *Contacts* -- refers to a subset of "exceptions" (above) in the pilot area which were identified by the TDUR Committees as **warranting** notification letters regarding potential inappropriate drug utilization (about 2,000 in the pilot counties) including those whose providers may not have been reached

- *Responses* -- refers to a subset of the "contact" group (above) who responded to the contact letter, or who appeared to have modified drug utilization patterns after the contact letter was issued, even though no written response was received (about 400 in the pilot counties).

It is important to note that Ernst & Young's direct benefit evaluation approach used only the "contact" and "response" groups. We did not conduct generalized statistical analysis using the "exceptions" level data. Use of the term "exceptions-based measures" relates specifically to the exceptions group of 17,000 (and its comparative cohort in the control group).

Response Group Alone

Measuring the benefits of the Response Group alone would entail measuring the health care utilization and costs for exceptions in the contact group whose re-review profiles showed a correction of the identified drug therapy problem. The before and after trends would be compared to determine if a benefit (reduction of costs) occurred after the intervention. No comparison group would be utilized. This alternative would not yield a result that could reliably be attributed to TDUR alone.

Response Group Versus Exceptions Not Contacted

The same measures would be applied as in the above alternative; however, similar measures would be obtained for the group of exceptions not selected by the

DURCs for provider notification. The advantages of this alternative are that (1) these two groups are separate and distinct, (2) the comparison group is similar to the control population in that no treatment occurred, and (3) the data is retrievable. The disadvantage is that the groups are made up of eligibles with dissimilar characteristics, by definition. One group was identified by the DURCs as having true drug utilization problems; the other was determined to represent problems with the TDUR computer system exceptions criteria. The service utilization and costs for health care of these two groups is likely to be different for reasons other than TDUR. Therefore, they are not comparable.

Response Group Versus Non-Response Group

Similar measures would be compared as above for the exceptions whose providers were notified and whose re-review profiles reflected a correction of the identified drug therapy problem versus exceptions whose providers were notified and whose re-review profiles reflected no correction of the identified drug therapy problem. The advantages of this alternative are that the Non-Response Group is most similar to the Response Group in that the same drug therapies were being utilized and same drug therapy problems were being experienced. However, it is not clear what methodology was used by the TDUR project manager and the DURCs to determine if a re-review profile showed change versus no change in the identified problem. Therefore, any methodology which utilizes this approach must address this limitation. We selected this alternative and developed specific review criteria to deal with the weakness cited above in our evaluation.

B. DESCRIPTION OF THE METHODOLOGY

Our methodology is principally organized into two critical activities. The objective of the first activity is to produce the quantitative data essential to measuring any improvement resulting from the TDUR program. This data would be analyzed and a determination made by a comparison of the response group to the non-response group. Any excess reduction in services or costs in the response group, beyond any general reduction in both groups as a whole, could be attributed to the TDUR program. This net reduction represents the maximum *direct* benefit derived from the TDUR program and can be compared to program costs to help determine cost-effectiveness. We were unable to measure indirect benefits from the pilot program in our study as no records exist to describe these.

The objective of the second activity is to augment the quantitative data by a review of program documentation, and interviews with staff of DHS, VCC, SRI, and members of the local DUR Committees. This qualitative information is important in determining how efficiently or effectively the program was implemented. It is conceivable that if the program were inefficiently or ineffectively implemented, that the measured benefits might be increased by improving the process.

The approach we selected for measuring the cost-effectiveness of TDUR in California includes measuring outcomes for the Response Group and comparing these to outcomes for the Non-Response Group. The data used in this research was drawn from the exceptions case profiles produced during TDUR pilot program operations. Case profiles included the exception profile, a re-review profile and any provider responses collected for each exception contacted. Appendix E contains a sample recipient profile used during this evaluation.

Four hundred Response Group cases, representing the universe of eligibles from whose data direct maximum benefit could be measured, and a random sample of 400 Non-Response Group cases were gathered from the TDUR pilot program operations location. Due to data inconsistencies in some of the cases, they were not usable, and we were left with 392 Response Group cases and 362 Non-Response Group cases. These cases represented the population of cases which were opened and reviewed during our study evaluation time frame. This timeframe included all cases for which reviews and re-reviews had been completed between July 1, 1989 and August 31, 1990. All of the cases were

subsequently closed by the DURCs and were assigned a status of either "P," reflecting provider therapy change, "U," reflecting patient utilization change, or "N," reflecting no change, at the time of closure.

Our methodology included developing independent criteria to conduct a double-blind physician analysis of these cases. The analysis was "blind" because the physician did not know how the case was analyzed during the DUR process. The purpose of the physician analysis was to provide an independent opinion regarding the status of each case when it was closed. The criteria and form used to record data for each case reviewed are included in Appendix F. The review procedure involved responding to a set of questions designed to determine if the identified drug therapy problem had been resolved and, if so, to determine if the resolution of the problem could be attributed to intervention by the DURCs.

Data from each of the 752 cases were entered onto a Recipient Data Base specifically designed for this evaluation. The kind of data captured included, for example, recipient identification, recipient age, recipient sex, drug therapy problem type, drugs used, number of physicians being used, and various other data.

We summarized prescription services, outpatient medical services, inpatient hospital admissions, and inpatient hospital days for a sample of contact group initial profiles and for all contact group re-review profiles. We then summarized prescription services, outpatient medical services, inpatient hospital admissions, and inpatient hospital days for all response group initial profiles, and for all response group re-review profiles. We measured the change in value of these indicators in both groups by comparing initial profile data to re-review profile data. We then compared the two groups to each other. As a result, we arrived at the net difference in service volumes between the non-response group and the response group. We determined that any positive difference reflected an improvement likely attributed to the TDUR System.

With respect to the pilot population of 300,000, we evaluated the total population who directly benefitted from the pilot; this represented only 392 recipients during the fourteen month period of July 1, 1989 to August 30, 1990 inclusive as identified by VCC. Though the possibility exists for second and third order benefits resulting from provider education and word of mouth, these do not represent direct benefits and are impossible to measure at the recipient history profile level of detail. If the direct benefits are used to

calculate average benefit per eligible, it is important to consider whether the total eligible population of 300,000 is used as the divisor, or whether the exposed eligible population of 90,000 is used as the divisor. Impact measured as an average reduction, whether in dollars or services, will of course, appear greater when considered over a smaller population.

C. SUMMARY OF RATIONALE FOR THE SELECTED APPROACH

The purpose of this subsection is to summarize key issues shaping the development of our evaluation approach. Deviations in the TDUR program implementation resulting in an absence of critical data have necessitated many elements of our approach, principally the conduct of independent research to determine the cost-effectiveness of the TDUR program. We believe our approach represents the best feasible alternative, given our limitations, and provides a reliable framework for achieving the objectives of this evaluation. A number of key issues were considered in developing this framework. These include:

- *Undertaking Independent Analysis of TDUR*
- *Use of Exceptions-Based Measures*
- *Elements of Our Approach Which Are Consistent With, or Similar to, the Original TDUR Analysis Approach.*

1. Undertaking An Independent Analysis of TDUR

Under normal circumstances we would not have undertaken an independent research project for this evaluation. We expected we would simply review the evaluative outputs of the DUR pilot project. Unfortunately, none of the products essential to carrying out this evaluation were produced as planned during the pilot. Also, we were told by DHS that a substantial part of the statistical reports generated to-date by VCC contained inaccurate or questionable data. Use of this general program data for evaluative purposes was discouraged by DHS.

Finally, SRI, under contract with VCC to conduct an evaluation of the TDUR program, planned delivery of its evaluative report to DHS sometime in early April, with a draft scheduled before that date. However, no firm date for either report was set by DHS. Our report to the State Legislature is required by May 1, 1991, and we could not wait for the pending SRI report.

Another option considered in lieu of an independent TDUR analysis was a review of the Department's **process** of implementing and evaluating the DUR pilot project. However, everyone involved in the project recognized that significant problems occurred in

the process of implementation. Original evaluation plans were scratched, data in reports were not reliable, progress reports were not issued, etc. We saw no practical benefit to the Legislature in a process review of the program.

For the reasons stated above, we concluded that a separate analysis of TDUR effectiveness was essential if we were to meet our contractual obligations to the Auditor General and the Legislature.

2. Use of Exceptions-Based Measures

Several difficulties arose in developing a suitable framework for undertaking an evaluative study which used "exceptions-based" measures. These include:

- ❑ Changing Exception Criteria Over Time
- ❑ Problems Defining the Exceptions Group
- ❑ Relevance to Public Policy Decisions.

Changing Exceptions Criteria Over Time

The maximum services component of the exceptions-level criteria were changed several times during the pilot study. In turn, this means that the exceptions population would be defined differently at different times, and that this would complicate using the exceptions-level information to define a reference population.

In other words, the continuing changes to maximum services exceptions-level criteria resulted in an exceptions population that was defined differently at different times. Furthermore, the structure of VCC's database did not support separating claims for people with maximum services exceptions from those of the rest of the exceptions population.

The original VCC study methodology prepared by Consolidated Consulting Group acknowledged this problem and included a means to correct for it. Also, this situation would be true in a real-life environment; i.e., criteria will be changed as conditions warrant.

Problems Defining the Exceptions Group

There are a number of issues which make the definition of the exceptions-level group difficult. One concern is that the subgroup must be defined independently of the treatment (something not done in this pilot program). This problem results in difficulties with comparing subgroups. This is an unavoidable problem, but it does not mean that the exceptions-level subgroups and data are unreliable for a comparative time-sensitive analysis. As mentioned earlier, it is not possible to achieve a truly experimental design in today's Medi-Cal environment. The TDUR pilot was never intended to be purely experimental, and it can never be so given the dynamic environment at play in Medi-Cal.

In our analysis, for example, we attempted to isolate the results of maximum services exceptions-level criteria from those of therapeutic exceptions-level criteria. We investigated whether these maximum services criteria changes had any real effect on decisions made by DUR committees to make written contact with providers, and we investigated whether maximum services criteria resulted in any direct program benefit.

The original proposal from VCC and CCG acknowledged this issue. The proposed analysis was characterized as quasi-experimental for just this reason. The original proposal went on to clarify the risks attributable to non-random assignment of recipients to subgroups, and to outline a complete methodology to isolate the effects of these problems from the outcome variables.

We believe that had the original methodology been followed, these problems could have been isolated and a reasonably reliable exceptions-level analysis completed. This is important, as only an exceptions or contact-level analysis can measure the **direct** effects of the TDUR program. VCC and CCG stated in their *DUR Program Evaluation, Detailed Technical Proposal* of March 15, 1989 that in order to substantiate that the indirect TDUR effect was caused by the TDUR contact, it is necessary to measure the direct effect. As VCC and CCG confirmed in their document, direct effect can only be measured at the treatment (contact) level.

Relevance to Public Policy Decisions

Using a subset of exceptions-level cases provides the only data available today to measure the direct impact of the TDUR program on the Medi-Cal community. Without these measures, there is no way to validate indirect benefit measured at the population-level and to reliably attribute it to the TDUR program. Therefore, our approach has considerable relevance to public policy decision-making.

3. Elements of Our Approach Which Are Consistent With or Similar to the Original TDUR Analysis Approach

The original TDUR analysis approach was described thoroughly in a proposal from CCG to DHS on March 15, 1989. This document described at length the complex needs of the analysis of TDUR. It effectively described why both an exceptions-level analysis and population-level analysis was necessary; the first would measure the direct impact of the TDUR program while the second would measure the indirect impact. The measure of the direct impact would be used to consider the reliability of the measure of the indirect impact.

The Department, VCC, and the new evaluation subcontractor, SRI, subsequently decided to abandon the exceptions-level measures and analysis. This was done for several reasons, including:

- Reporting and data problems at the exceptions-level
- Absence of sufficient control group data for exceptions-level comparisons
- Arguments with the exceptions-level measurement concepts.

As a result, the Department is only analyzing population-level data in determining the impact of the TDUR program. Our methodology included a cohort analysis of exceptional recipients in two groups: the contact, "positive response" group, and the contact "no response" group. Both are subgroups of the exceptions group.

We believe that unless there was a measurable, direct impact from the TDUR program, any reported indirect impact remains unverified. Due to deviations in the TDUR implementation, the only way to measure direct impact was to examine individual exception review/re-review profiles for the contact group, as we have done.

D. CONCLUSIONS

Performing an analysis at the exceptions-level is more complex and requires accommodation of a larger number of variables than doing only a population-level analysis. On the other hand, as discussed in the original VCC/CCG proposal, some form of exceptions-level analysis had to be completed in addition to population-level analysis for the study findings to carry reliability as far as benefit attributable to the TDUR program.

VCC/CCG offered a methodology which addressed the concerns raised in doing exceptions-level analysis. As it turned out, the demise of CCG and the extended amount of time taken to replace that firm made it impossible for SRI to complete all the tasks outlined in the VCC/CCG proposal. Also, the Department's position regarding the questionable reliability of VCC reporting contributed to a decision to abandon all but the population-level analysis. We wish to point out that problems attributable to contact or exceptions-level measures roll-up to the population-level. To ignore these issues at the subgroup level only serves to ignore the most fundamental question in this pilot: Is there any direct impact from the TDUR pilot program?

Ernst & Young elected to proceed with an analysis of the "contact" and "response" subgroups of the "exception" group because we believed it was important to make an effort to assess the direct impact of the TDUR program on the pilot Medi-Cal community.

APPENDIX E

SAMPLE PROFILE EXCEPTION REPORT

ERNST & YOUNG

REPORT NO. DU-O-81
DATE: 07/02/90 TIME: 22:45:05
CYCLE ENDING:05/31/90

STATE OF CALIFORNIA
DRUG UTILIZATION REVIEW SYSTEM
EXCEPTION PROFILE REPORT

COUNTY: 50 STANISLAUS

RECIPIENT ID: AGE: 54 SEX: M

CURRENT MONTH EXCEPTIONS

LINE NO.	REF NO.	DESCRIPTION	THERAPEUTIC CRITERIA
1	08331	DRUG / DRUG	NORTRIPTYLINE-MAY DECREASE EFFICACY OF PHENYTOIN

DRUG / DIAGNOSIS HISTORY (3 MONTHS)

DATE OF SERVICE	DRUG / TC / GENERIC CODE		DIAGNOSIS CODE	DIAGNOSIS DESCRIPTION	QTY DISP	DOSAGE FORM	STRENGTH/ PRESC #	PROVIDERS PHYSICIAN	PROVIDERS PHRM/HOSP	INPAT LOS (Y/N)
03/13/90	SU4300D	ALTERNAGEL			440	LIQUID	120.000	00A3[redacted]	PHA[redacted]	
	01 / 08040	ALUMINUM HYDROXIDE					209121			
03/13/90	MR1764A	CARAFATE			100	TABLET	1.000	00A3[redacted]	PHA[redacted]	
	01 / 08200	SUCRALFATE					206429			
03/15/90			3681	SUBJECTVE VISUAL DISTU					YYY[redacted]	
03/23/90	SU4300D	ALTERNAGEL			440	LIQUID	120.000	00A3[redacted]	PHA[redacted]	
	01 / 08040	ALUMINUM HYDROXIDE					209121			
03/23/90	PK1770G	METOCLOPRAMIDE HCL			100	TABLET	10.000	00A[redacted]	PHA[redacted]	
	99 / 21020	METOCLOPRAMIDE HYDROCH					207079			
03/27/90	PD2407L	DILANTIN			100	CAPSULE	100.000	00C[redacted]	PHA[redacted]	
	48 / 17700	PHENYTOIN SODIUM EXTEN					204241			
04/03/90	MR1764A	CARAFATE			100	TABLET	1.000	00A[redacted]	PHA[redacted]	
	01 / 08200	SUCRALFATE					206429			
04/13/90	SA2619F	PAMELOR			100	CAPSULE	50.000	00A[redacted]	PHA[redacted]	
	11 / 15283	NORTRIPTYLINE HYDROCHL					214428			
04/17/90	SU4300D	ALTERNAGEL			440	LIQUID	120.000	00A[redacted]	PHA[redacted]	
	01 / 08040	ALUMINUM HYDROXIDE					209121			
04/25/90	NY3110C	PROPRANOLOL HCL			120	TABLET	40.000	00C[redacted]	PHA[redacted]	
	76 / 20632	PROPRANOLOL HYDROCHLOR					199694			
04/27/90	MR1764A	CARAFATE			100	TABLET	1.000	00A[redacted]	PHA[redacted]	
	01 / 08200	SUCRALFATE					206429			
04/27/90	SU4300D	ALTERNAGEL			440	LIQUID	120.000	00A[redacted]	PHA[redacted]	
	01 / 08040	ALUMINUM HYDROXIDE					209121			
04/30/90	PD2407L	DILANTIN			100	CAPSULE	100.000	00C[redacted]	PHA[redacted]	
	48 / 17700	PHENYTOIN SODIUM EXTEN					204241			
05/11/90	SU4300D	ALTERNAGEL			440	LIQUID	120.000	00A[redacted]	PHA[redacted]	
	01 / 08040	ALUMINUM HYDROXIDE					209121			
05/11/90	PK1770G	METOCLOPRAMIDE HCL			100	TABLET	10.000	00A[redacted]	PHA[redacted]	
	99 / 21020	METOCLOPRAMIDE HYDROCH					207079			

REPORT NO. DU-O-81
DATE: 10/24/90 TIME: 22:04:41
CYCLE ENDING:09/30/90

STATE OF CALIFORNIA
DRUG UTILIZATION REVIEW SYSTEM
EXCEPTION PROFILE REPORT

COUNTY: 50 STANISLAUS

RECIPIENT ID: [redacted] AGE: 54 SEX: M

** PROFILE FOR RE-REVIEW BY LOCAL DURC **

EXCEPTION HISTORY

DATE	REF NO.	PROBLEM	THERAPEUTIC CRITERIA
05/90	08331	DRUG / DRUG	NORTRIPTYLINE-MAY DECREASE EFFICACY OF PHENYTOIN

DRUG / DIAGNOSIS HISTORY (3 MONTHS)

DATE OF SERVICE	DRUG / TC / GENERIC CODE		DIAGNOSIS CODE	DESCRIPTION	QTY DISP	DOSAGE FORM	STRENGTH/ PRESC #	PROVIDERS PHYSICIAN	PROVIDERS PHRM/HOSP	INPAT LOS (Y/N)
07/10/90	MR1764A	CARAFATE			100	TABLET	1.000	00A[redacted]	PHA[redacted]	
	01 / 08200	SUCRALFATE					206429			
07/10/90	SU4300D	ALTERNAGEL			440	LIQUID	120.000	00A[redacted]	PHA[redacted]	
	01 / 08040	ALUMINUM HYDROXIDE					209121			
07/20/90	PD2407L	DILANTIN			100	CAPSULE	100.000	00C[redacted]	PHA[redacted]	
	48 / 17700	PHENYTOIN SODIUM EXTEN					204241			
07/24/90	SU4300D	ALTERNAGEL			440	LIQUID	120.000	00A[redacted]	PHA[redacted]	
	01 / 08040	ALUMINUM HYDROXIDE					209121			
07/30/90	SA2619F	PAMELOR			100	CAPSULE	50.000	00A[redacted]	PHA[redacted]	
	11 / 15283	NORTRIPTYLINE HYDROCHL					214428			
08/03/90	MR1764A	CARAFATE			100	TABLET	1.000	00A[redacted]	PHA[redacted]	
	01 / 08200	SUCRALFATE					206429			
08/03/90	SU4300D	ALTERNAGEL			440	LIQUID	120.000	00A[redacted]	PHA[redacted]	
	01 / 08040	ALUMINUM HYDROXIDE					209121			
08/08/90	NY3110C	PROPRANOLOL HCL			120	TABLET	40.000	00C[redacted]	PHA[redacted]	
	76 / 20632	PROPRANOLOL HYDROCHLOR					217772			
08/17/90	PK1770G	METOCLOPRAMIDE HCL			100	TABLET	10.000	00A[redacted]	PHA[redacted]	
	99 / 21020	METOCLOPRAMIDE HYDROCH					207079			
08/27/90	MR1764A	CARAFATE			100	TABLET	1.000	00A[redacted]	PHA[redacted]	
	01 / 08200	SUCRALFATE					206429			
08/29/90	PD2407L	DILANTIN			100	CAPSULE	100.000	00C[redacted]	PHA[redacted]	
	48 / 17700	PHENYTOIN SODIUM EXTEN					204241			
09/07/90	SU4300D	ALTERNAGEL			440	LIQUID	120.000	00A[redacted]	PHA[redacted]	
	01 / 08040	ALUMINUM HYDROXIDE					209121			

JULY 31, 1990

MODESTO, CA 95350

RE: RAY L GOODSON, MEDICAID ID: 525703538

1. HAS THE INFORMATION CONCERNING THIS PATIENT BEEN USEFUL TO YOU?
YES ___ NO ✓

COMMENTS:

2. WOULD THIS KIND OF INFORMATION GENERALLY BE USEFUL TO YOU?
YES ✓ NO ___

COMMENTS: LONG'S DRUG STORE AUTOMATICALLY CHECKS FOR ALL DRUG INTERACTIONS BY COMPUTER

3. AS A RESULT OF THIS INFORMATION, DO YOU ANTICIPATE A POSSIBLE CHANGE IN DRUG THERAPY? YES___ NO✓

IF SO, WOULD YOU PLEASE TELL US WHAT YOU ANTICIPATE:

4. AS A RESULT OF THIS INFORMATION, DO YOU ANTICIPATE CLOSER MONITORING AND/OR COUNSELING OF THIS PATIENT? YES ___ NO ✓

COMMENTS: Will continue routine dilantin levels -- last seizure was in 1984 -- pt is doing quite well on current treatment plan

5. DO YOU HAVE ANY COMMENTS OR SUGGESTIONS ABOUT THIS PROGRAM OR HOW WE MAY BETTER ASSIST YOU WITH THE CONTINUING CARE OF YOUR PATIENTS?

COMMENTS:

This is the first report I have received from you. I would need to see 8-10 more before I could give an opinion as to how helpful it would be.

JULY 31, 1990

RE: MEDICAID ID:

1. HAS THE INFORMATION CONCERNING THIS PATIENT BEEN USEFUL TO YOU?
YES ✓ NO ___

COMMENTS:

2. WOULD THIS KIND OF INFORMATION GENERALLY BE USEFUL TO YOU?
YES ✓ NO ___

COMMENTS:

3. AS A RESULT OF THIS INFORMATION, DO YOU ANTICIPATE A POSSIBLE CHANGE IN DRUG THERAPY? YES___ NO✓

IF SO, WOULD YOU PLEASE TELL US WHAT YOU ANTICIPATE:

4. AS A RESULT OF THIS INFORMATION, DO YOU ANTICIPATE CLOSER MONITORING AND/OR COUNSELING OF THIS PATIENT? YES ✓ NO ___

COMMENTS:

5. DO YOU HAVE ANY COMMENTS OR SUGGESTIONS ABOUT THIS PROGRAM OR HOW WE MAY BETTER ASSIST YOU WITH THE CONTINUING CARE OF YOUR PATIENTS?

COMMENTS: NONE

APPENDIX F

SAMPLE PHYSICIAN ANALYSIS RECORD

ERNST & YOUNG

DRUG UTILIZATION REVIEW

PHYSICIAN ANALYSIS RECORD

RECIPIENT ID _____ SEQUENCE NUMBER 2

REVIEW CYCLE END DATE 5/31/89 RE-REVIEW DATE 9/30/89

REPORT PRINT DATE 8/11/90

DESCRIPTION OF PROBLEM ~~multiple docs for controlled substance~~ drug abuse

I. NATURE OF CHANGE:

1 QUANTITY __ INCREASE X DECREASE
__ DRUG __ TYPE __ NAME
1 NUMBER OF DOCTORS PRESCRIBING __ INCREASE X DECREASE
__ PRESCRIPTION STRENGTH __ INCREASE __ DECREASE
__ FREQUENCY OF REFILLS __ INCREASE __ DECREASE
__ PROBLEM CORRECTED (WEIGHT 5)
~~2~~ TOTAL

II. BEHAVIOR CHANGE:

CHANGE IN PATIENTS BEHAVIOR
1 -1 0 1 2 3 (0 = NO CHANGE)

CHANGE IN DOCTORS BEHAVIOR
3 -1 0 1 2 3 (0 = NO CHANGE) __ IRRITATION
-1 LETTER RECEIVED > 2 MONTHS INTO RE-REVIEW CYCLE

CHANGE IN PHARMACISTS BEHAVIOR
__ -1 0 1 2 3 (0 = NO CHANGE) __ IRRITATION
__ LETTER RECEIVED > 2 MONTHS INTO RE-REVIEW CYCLE

3 TOTAL

III. EFFECT OF LETTER:

FROM DOCTOR COMMENTS, IS CHANGE A RESULT OF T-DUR LETTER?
3 -1 0 1 2 3 (0 = NO RELATION)

FROM PHARMACIST COMMENTS, IS CHANGE A RESULT OF T-DUR LETTER?
__ -1 0 1 2 3 (0 = NO RELATION)

3 TOTAL

IV. COMMENTS Through letter late in cycle, abuse tapered off shortly after letter

cc: Members of the Legislature
Office of the Governor
Office of the Lieutenant Governor
State Controller
Legislative Analyst
Assembly Office of Research
Senate Office of Research
Assembly Majority/Minority Consultants
Senate Majority/Minority Consultants
Capitol Press Corps