



USAID
FROM THE AMERICAN PEOPLE

BULGARIA PHARMACEUTICAL STUDY

METHODOLOGY

AUGUST 2005

This publication was produced for review by the United States Agency for International Development. It was prepared by the IRIS Center, at the University of Maryland.

BULGARIA PHARMACEUTICAL STUDY

METHODOLOGY

DISCLAIMER

The author's views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

SUBMITTED BY

The IRIS Center, University Research Corporation International

SUBMITTED TO

David Meyer, Cognizant Technical Officer

CONTRACT NUMBER

PCE-I-00-97-00042-00, Task Order 8

AUTHORS

Omar Azfar, Patrick Meagher, Diana Rutherford

The Bulgaria study referred to in this report is: Meagher, P., Azfar, & Rutherford, D. (2005). *Governance in Bulgaria's Pharmaceutical System: A Synthesis of Research Findings*. College Park: IRIS Center, University of Maryland.

DOWNLOAD

www.irisprojects.umd.edu/anticorruption

CONTACT INFORMATION

IRIS Center
University of Maryland
Department of Economics
2105 Morrill Hall
College Park, MD 20742, USA

E-mail: info@iris.econ.umd.edu

Phone: +1.301.405.3110

Fax: +1.301.405.3020

Web: www.iris.econ.edu

Introduction

This study is part of a larger project, **testing program-relevant corruption assessment methodologies**. The USAID E&E Bureau asked the IRIS Center to develop a methodology to assess corruption in selected sectors, particularly at the level of micro-organizations (such as commercial courts, tax administration offices, or schools). The resulting studies incorporate the conceptual framework developed by USAID/EE, focusing on the role of Transparency, Accountability, Prevention, Enforcement and Education (TAPEE) as institutional requirements of integrity – i.e., efficient and effective governance free of corruption.

The study in Bulgaria focused on the processes used in the selection and procurement of pharmaceuticals used in the healthcare system. The research had two components. The **drug selection component** focuses primarily on the two major selection processes: the Positive Drug List (PDL) and the National Health Insurance Fund (NHIF) Reimbursement List. We also reviewed, more briefly, the selection process involved in the Ministry of Health (MOH) Expensive Drugs List. The **procurement component** deals with the purchase of drugs by hospitals. The medicines procured are listed on hospital formularies, which must in turn be selected from the Positive Drug List (unless the hospital operates outside the national health insurance system).

In the present report, we explain the methodologies used in the Bulgaria study, taking each of the two components in turn. We highlight lessons learned in our use of the methodologies.

Drug Selection for Central Lists

In this research, we used qualitative methods, i.e. interviews, review of official documents and procedures, analysis of media reports, and comparisons of processes and results to international standards. Our approach emphasized structured interviews, based on detailed protocols, with key officials from the relevant government departments, representatives of the pharmaceutical industry (foreign and domestic), and independent experts. The interviews included initial “key informant” interviews for background, as well as some 30 structured interviews with officials and firms.

This methodological choice was dictated by the nature of the central listing processes. They are few in number and involve a finite group of officials, experts, and drug companies. Thus, no statistically valid survey was feasible, and a flexible approach was necessary in order to collect information from persons who might not agree to respond to a questionnaire. There are further complications. In the selection processes, it is a question of high-level or “grand” corruption. There would be serious legal, political, and personal consequences for anyone implicated (as contrasted with low-level bribery, which is often tolerated). Thus, reticence is a serious concern that we tried to mitigate in the interviews (mainly through the ordering and wording of questions).

Initial background research

We began the research (on both components) with the collection of background information and an intensive two weeks of key-informant interviews in Bulgaria. These interviews were greatly assisted by Bulgarian staff of the USAID mission (in the DG and Health sectors), other USAID contractors, and a Sofia-based research organization (Vitosha Research, affiliated with Coalition 2000). At this point, our inquiry was broadly concerned with pharmaceuticals in the Bulgarian healthcare system, but the decision had not yet been made whether to focus on central selection, hospital-based procurement, distribution via pharmacies, or some combination. We therefore interviewed a broad cross-section of aid donors, contractors, central government officials (primarily in the Ministry of Health and NHIF), local government officials, pharmaceutical companies and industry associations, hospital personnel, health NGOs and research institutes, and members of the physicians' professional association. We also collected numerous reports and official documents.

On the strength of the above material, we produced an inception report laying out the key governance issues in the pharmaceutical system, and presenting options to USAID for the research design. After some discussion, it was mutually decided by USAID and IRIS that the full study would have one component each dealing with the selection and procurement processes (the second component is discussed in Part 3 below). For both components of the study, we created a team that included IRIS researchers as well as external experts on healthcare administration – including Jillian Cohen and Judith Fisher of the University of Toronto, and the International Healthcare and Health Insurance Institute (IHII) based in Sofia.

Institutional integrity assessment

For the next phase of the research, the IRIS team and its external collaborators developed and applied a protocol of research questions, which were grouped according to the relevant integrity factor in the USAID/EE TAPEE framework. A summary of that protocol is presented in Annex 1. These questions, with some specific adaptations according to context, were used in interviews and as a guide to the collection of documents and data.

In this phase, we conducted 30 interviews, including six with members of the Commission on the Positive Drug List, six with Ministry of Health officials, four with manufacturers, two with NHIF officials, two with members of parliament, and ten with health NGOs. Of these, some two-thirds were structured in accordance with the protocol, while the rest had to be done more informally, based loosely on the protocol.

We supplemented the interviews with documentary research. We formulated and submitted official information requests under the Access to Public Information Act

(APIA). IHHI took the lead role here. Some 15 requests were made to the Council of Ministers, the Ministry of Health, and NHIF. All requests received responses, although a few of the responses were partial and would have required more follow-up than we had time for. The material gathered this way included regulations and orders, minutes of meetings, lists of commission members, and other documents. In addition to providing some necessary information, the APIA process itself served as a test of the system's transparency. Further, we reviewed additional background material bearing on the integrity of the selection system – including professional ethics norms, public official ethics rules, and government-wide integrity and anti-corruption systems – and assessed their bearing on the selection processes.

We used the above data to assess the selection system's integrity (and, conversely, its vulnerability to corruption). We used as our integrity standards the factors presented in USAID/E&E's analytical framework, TAPEE – transparency, accountability, prevention, enforcement, and education. We applied these factors to our findings on the drug selection processes, mainly the PDL and the Reimbursement List, with an additional brief look at the Expensive Drugs List. We addressed specific questions regarding TAPEE factors and rated the results on a four-point scale (poor, average, good, excellent), based on comparisons of Bulgaria's system to “best” practice as set forth in international standards and the analytical literature.¹ These ratings are intended to be informative, indicating the likelihood of corruption, without necessarily being conclusive. The combined list of questions concerning all these factors, along with the best practice benchmarks, appears in Annex 1.

Assessment of political-economic drivers and outcomes

USAID/E&E's TAPEE framework posits a causal nexus between institutional integrity and corruption. All other things equal, one would anticipate an inverse relationship, i.e. higher integrity results in lower corruption.

In the context of the drug selection process, as we suggested above, a statistical test of causation was not possible. The selection processes involve three groups of decision-makers, and overall a relatively small number people. We re-stated the causal hypothesis as follows:

1. Political-economic factors, or corruption “drivers,” e.g.:
 - Drug firms' scramble to grab shares in a market dominated by government
 - Politicians, supported by industrial and patients' lobbies, seeking gain (political and personal) by influencing market share (selections)

→ Lead to:

2. Corruption, e.g. market allocation based on bribes, favors, illicit relationships – as evidenced by:
 - Experiences, perceptions, reports, and cases of corruption

¹ This includes MSH & WHO (1997), and Cohen, Cercone and Macaya (2002).

- Distortions in drug selection and reimbursement price-setting

→ If and only if there are:

3. Opportunities presented by weaknesses in government institutional integrity –
 - As measured by TAPEE.

Our research on corruption “drivers” in the Bulgarian political economy was also qualitative, due to time and resource constraints (as well as our judgment that quantitative methods would not be cost-effective). We relied partly on findings from our structured interviews and to a lesser extent on findings from the hospital procurement survey (i.e. responses to questions about aggressive marketing by drug firms, see below). We also used background research on Bulgaria and its pharmaceutical market, as well as comparative studies of market pressures and outcomes in other countries’ drug approval and selection processes – including the U.S. and developing countries. In addition, we reviewed official documents concerning drug company applications for listing, as well as parliamentary testimony and legal filings by the pharmaceutical industry associations.

Our key source of information on drivers was a “media analysis” conducted by IHIII. This involved a review of Bulgarian media reports on drug selection and marketing practices during the 12 months from mid-2003 to mid-2004, with a smaller follow-up review covering the rest of 2004. This review covered more than 5,000 media reports, which included coverage of several scandals in the pharmaceutical system. Our analysis of this material enabled us to assess the pressures placed on the selection system – influences that may drive corrupt practices.

As for the outcomes of the selection process, we were concerned with two dimensions: (i) the procedural regularity of the process in practice, including allegations and evidence of corruption in the system; and (ii) the technical quality and cost-effectiveness of choices. On the first point, we sifted through the results of our background research, the structured interviews, and the media analysis for evidence – of both procedural compliance and non-compliance, as well as the presence of corruption (weighing allegations and reports in terms of their credibility).

On the second point, we looked at the technical soundness of drug choices and the prices set for NHIF reimbursement. The technique here was to compare these outcomes to international benchmarks, with the objective of identifying anomalies that would raise “red flags” indicating the probable presence of corruption. Thus, a pharmacy expert on our team reviewed the PDL and NHIF Reimbursement List, comparing the selected compounds and brand-names against WHO drug selection standards and its List of Essential Drugs. She found a number of anomalies that seemed to fly in the face of international best practice – i.e. choices that seemed highly unsuitable on medical and cost grounds and that had no credible justification.

Regarding costs, IHIII created a sample of 20 drug compounds for international comparison in both the selection and procurement parts of the study – of these, seven

appeared on the NHIF List. They then chose a set of five transition countries from among the comparator countries used by the Bulgarian NHIF to set reimbursement price guidelines. All of these countries had some brand-name products in common with those that appear within the seven international non-patent name categories on the NHIF List. We compared prices, again looking for major anomalies that would suggest probable corruption.

This analysis of outcomes was not meant to produce legally-sufficient evidence of corruption, nor to level accusations against any particular person or agency. Rather, it was meant as a way to assess patterns in systemic outcomes that point to possible corruption. Importantly, we did not equate minor differences or lapses in professional judgment with corruption. Rather, we identified results that reached a threshold of significance as serious anomalies – like statistical outliers – interpreting these as probable outcomes of corruption or other systemic failures. Triangulating these findings on outcomes with our results on drivers and integrity factors gave us a holistic picture of the quality of governance in the system. This picture did not provide conclusive proof of corruption, but it did identify areas that give cause for serious concern, and this led to well-grounded recommendations about steps to be taken that would constrain the corruption that does appear to exist in the system.

Drug Procurements by Hospitals

In the study of Bulgarian hospitals, we decided to conduct a survey based study to collect representative data on the hospital procurement system. We worked with local partners – the International Healthcare and Health Insurance Institute (IHII) and its survey affiliate, FACT Marketing – to carry out a survey of 148 hospitals (out of 236 medical institutions in Bulgaria) that agreed to participate in the research. IN addition to the suppliers of medicines to hospitals, respondents included doctors, nurses, pharmacists, evaluation committee members, and hospital directors. We used the survey results, with corroborating evidence, to detect significant levels of corruption, to assign values to the TAPEE factors, and to assess any relationship between these factors and corruption.

The main strength of survey research, which asks a large number of respondents exactly the same question with identical lists of answer options, is that it produces answers which are comparable and easily presentable. But collecting useful data on a complex process like hospital procurement in this way is a challenging task. The analyst must identify specific scenarios that are comparable across hospitals. This requires detailed knowledge of the procurement process. We acquired the necessary knowledge by a combination of document review, key informant interviews, focus groups, and pre-tests: We were assisted in all these efforts by our Bulgarian sub-contractor IHII, a firm that specializes in health care research and policy advice. These pre-survey steps are described next.

Pre-Survey Investigations

Document review

The first phase of the study involved a review of documents of the Bulgarian pharmaceutical system. This included both the laws themselves and reports about the sector. Some documents were only available in Bulgarian: These were translated or summarized by IHHI. These documents, in particular the law on procurement, were vital in the design of the survey instruments, which followed the various steps of the procurement process.

Key Informant interviews

The document review served only as a starting point for the survey design. The next step was a combination of key informant interviews and focus groups, to collect information on de facto processes that were closely relevant to our study.

In a key informant interview, the analyst interviews a participant or an informed observer of the procurement process. In the hospital procurement study, we interviewed hospital directors, pharmacists, and evaluation committee members, and people who oversee the hospital at the ministries of health or finance, or audit agencies. Ideally, such an interview is conducted in private, unless a translator is needed. For our study, interviews were conducted both by an ex-pat Jillian Cohen, with a translator present and by IHHI staff where no translator was needed.

Focus Groups

In focus groups, a moderator gathers several participants into a room, and leads a focused discussion among them. Focus groups have the advantage of gathering information from several participants at once. Another advantage of focus groups is the likelihood of participants raising previously unidentified issues about which other participants may contribute their experience and knowledge. Another potential advantage is that sometimes the candor of the participants increases as they feed off each others reports and complaints. We conducted focus groups with doctors, nurses, pharmacists and patients. It was not possible to conduct focus groups with hospital directors.

Pre-tests

Once the instruments are designed it is important that they be pre-tested before the survey is carried out. This is important because it helps ascertain whether the survey is of a feasible length, and whether the questions are understood by the respondents. Additionally, the analyst can follow the survey with a question and answer session on interesting issues that were raised, and ask whether important factors have been omitted. We did in fact conduct these pre-tests, and learnt about some important questions to add to the questionnaire. We were also reassured that the surveys had in fact covered most important issues.

Methodological lessons from pre-survey investigations

The information collected in the pre-survey process was of great importance. What follows is a partial list of lessons we learned in the pre-survey phase that informed the design of the survey instruments.

We learnt important lessons about sampling: Patients would not have useful information, and hence should not be sampled. And, some hospitals do not conduct their own procurement and therefore should be omitted from the sample.

We learnt important lessons on issues to exclude: Drug availability and quality did not appear to be important issues.

We learnt important lessons about issues to include: For instance, one practice we learned about was that a supplier may deliberately offer an unrealistically low bid, and be selected by a hospital, on the tacit agreement between the supplier and hospital director that the price would be raised later by a “contract annex” (amendment). Consequently we asked about this practice and found that contracts are amended in a number of hospitals.

The methodological implications that follow from this are that pre-survey investigations are vital. Without them we would end up choosing the wrong samples and asking the wrong questions. Indeed, each stage of pre-survey investigation is invaluable: important lessons for survey design and sampling were learnt in each of the document reviews, key informant interviews, focus groups and pre-tests. Analysts should plan to conduct each of these activities and garner the resources to conduct them well.

Survey

Respondent Selection

Pharmaceutical procurement is a complicated process. Some information is only possessed by some participants, and hence a number of respondents have to be surveyed to get a full picture of the actual process. Surveying multiple respondents was also useful for pieces of information that several respondents possess, because this allows cross-checks on the data. A total of 5 survey instruments were designed and implemented on hospital staff: these were hospital directors, pharmacists, evaluation committee members, doctors and nurses. In addition we tried to survey suppliers, but were largely unsuccessful.

Survey method.

We used both face-to-face surveys, and self-administered surveys in this study. The surveys for pharmacists, evaluation committee members, and the first half of the director’s survey were administered by interviewers (face-to-face). Doctors and nurses were surveyed using a self-administered instrument. The second half of the director’s

instrument was also self administered. The reasons for choosing self administration for the doctors and nurses surveys was that we were planning to use these surveys to gather information on corruption. We used self-administered surveys on the presumption that they lead to greater candor. A further advantage was that self administration reduces costs. The second half of the director's survey was self administered because it asked about detailed information on prices and quantities etc, information that must be gathered from files. In fact, this turned out not to work so well, because some directors appeared to misreport units of medicine where we wanted to compare price per unit on the same medicines across hospitals. Because of self administration this couldn't easily be checked.

Information on various aspects of integrity was collected from directors, pharmacists, and evaluation committee members. This information was collected using both questions on what actually happens (e.g., How many bidders were ranked) and questions on what would happen in hypothetical situations (What would happen if an evaluation committee member was caught taking a bribe). For things that have actually happened everywhere it is best to ask about what happened. But several aspects of integrity are inherently conjectural. "What happened when the evaluation committee member was caught taking a bribe?" will only produce comparable data if some evaluation committee member has actually been caught taking a bribe in each hospital. Otherwise there will be lots of missing data. Hence, in such a situation it is best to ask a conjectural question, which is what we did. Results are presented in the mission report.

Reticence

One important innovation we used in the Bulgaria methodology report was the "Envelope method for identifying reticent respondents." This method, involves two envelopes, each of which has two questions in it. The method is implemented as follows (sample instruments are provided in Annex 2):

The respondent is handed the first envelope, and told to answer the underlined question, without revealing which question is underlined. The respondent is told not to make any gestures which may reveal which question is asked. The envelope contains the following two questions

Have you ever stolen money or medicine?
Is $3+3 = 6$?

The first question is underlined in 90% of the cases, and the second question in 10% of the cases. If the respondent says anything other than yes or no, or makes any gesture that may indicate which question was asked, the interviewer asks the respondent to refrain from doing so when the second envelope question is asked.

The respondent is then handed the second envelope. This envelope contains two questions. The second, innocuous, proverbially simple, mathematical, question is always underlined.

Did you ever accept a bribe?

Is $2+2 = 4$?

Since only the mathematical question is underlined we know that any respondent answering no must be reticent, in the sense that they are not willing to give a candid answer to a sensitive question.

In the self administered version, the respondents are told that they should simply answer the questions, and not tell anyone which question they answered.

The method identifies around 1/3 of the respondent as being reticent, with the level of reticence varying across respondent categories from 28% (doctors) to 36% (nurses). There is no perceptible pattern of reticence varying by either respondent rank or survey method. The percentages for each category are shown in Table 1. This finding, that 1/3 of the sample appear to not be willing to answer sensitive questions, is of great methodological importance. Conversely, so is the finding that 2/3 of the sample does appear to answer questions about corruption in a candid fashion. It implies that data on corruption in Bulgarian hospitals is flawed, but not fatally flawed. In addition, it suggests a way to improve the quality of data: the reticent respondents can be removed and the analysis can proceed with the non-reticent respondents. This is what we did.

There is a significant pattern within categories by age of respondent. This difference is clearly significant at 1% in each of the three large samples (nurses, doctors, and evaluation committee members), and for the pooled sample. This is similar to the results found in neighboring Romania (Azfar and Murrell 2005). One possible explanation is that the number of years spent under communist rule may have influenced respondents' reticence.

Does the method actually identify reticent respondents, in the conventional sense that the respondents are reticent in their answers to sensitive survey questions asked in a standard way? We attempt to answer this question by examining whether there are differences in the way those who said no to second envelope question, answer "none" to the following question:

Doctors in many countries require informal payments before they will treat patients. How many do you think doctors in Bulgaria demand informal payments from patients before treating them?

Answers:

1. None 2. Very few 3. Some 4. Most of them 5. Almost all 6. All

We chose this question and answer, to examine the validity of the envelope method, for two reasons: The question is not about the respondent's own behavior, and hence the problem of the potential correlation of guilt and reticence is finessed. The extent of informal payments, and the existence of several media reports on the practice, make it implausible that any health care professional could have truthfully given this answer.

Distinctions between (say) “very rarely” and “most” may on the contrary depend on the respondent’s experience, which like guilt can be correlated with reticence. In fact, the differences between the reticent and non-reticent sub-samples are less significant, and sometimes even of the wrong sign for questions about ones own behavior. Disentangling guilt and reticence would be a worthwhile but challenging task for future work.

Table 1. Numbers of Respondents (out of 148 hospitals studied) ²

	Hospital Directors	Pharmacists	Evaluation Committee Members	Doctors	Nurses	All
Total # of respondents	139	111	440	551	707	1952
Non-reticent (#)	90	73	317	398	455	1297
Reticent (% of total)	35.3%	34.2%	28.0%	27.8%	35.6%	33.5%
Do doctors demand informal payment? Percentage who say none in reticent and candid categories.						
Proportion of reticent saying “none”	11.7	12.8	8.8	10.9	9.7	10.2
Proportion of candid saying “none”	7.9	6.9	5.6	7.9	6.2	6.8
t-stat	0.73	1.02	1.20	1.12	1.72*	2.66** *

The results show that respondents identified as reticent by the envelope question, are substantially more likely to say “none” in their response to the question about Bulgarian doctors taking informal payments. For each category of respondent, reticent respondents are about one and a half times more likely to say “none” than are candid respondents. The differences within any category of respondent are not significant, but if the data is pooled the difference is significant at 1%.

Internal Cross Checks

The significant proportion of reticent respondents was eliminated from the data. In our analysis linking corruption to integrity we used data from doctors and nurses for the corruption variable, and data from the evaluation committee members for the integrity variables. We chose evaluation committee members rather than the perhaps better

² Very few of the suppliers returned our survey forms. None of the suppliers answered the envelope questions, so we have no reticent/non-reticent division. It does however, seem that all of the suppliers are reticent in the conventional sense, given their refusal to answer many questions, and the small number of suppliers who were willing to answer even some of the questions.

informed pharmacists and directors for two reasons: both having to do with the fact that there were multiple evaluation committee members surveyed at each hospital. First, the elimination of reticent directors and pharmacists would have led to missing data on the hospital, while evaluation committee members could be dropped without losing information on the hospital. Second, the existence of multiple evaluation committee members at each hospital meant that consistency checks could be conducted on the data.

After eliminating reticent respondents a number of consistency checks were conducted on the data. We call this process an internal cross check because it is checking survey data against survey data. It is not internal in the narrow sense of checking a respondent's answers against herself, but rather in the sense that it is checking survey data against survey data.

The basic logic of a consistency check is that if two respondents are reporting on the same event they should give the same answer, or at least similar answers. For some questions the expectation is to have a close fit and a large proportion of respondents actually agreeing on the answer. For instance, we should expect identical answers from different evaluation committee members to "how many suppliers were ranked by the committee". For other questions, which are either hypothetical, or have imprecise answer scales, we would expect less perfect agreement. In general we did find agreement between what respondents said. The agreement was closer for questions where we would expect closer agreement.

For the corruption variable, the data passed the consistency check in the sense that nurses and doctors were more likely to report there was corruption if other nurses and doctors reported there was corruption. But the correlation was far from perfect. This is to be expected, both because the answer scale was imprecise, and because in this instance reticence may also be clouding the answers. We had eliminated respondents identified as reticent by the envelope questions, but the method may not have eliminated all reticent respondents.

External Cross Checks

Internal cross checks are only partially reassuring because correlated errors can also lead to "consistency". Added credibility can be gained by checking survey data against data collected by some other means. We tried two methods: price and quantity analysis, where prices or quantities that were too high would be regarded as indications of corruption, and an analysis of audit reports.

One way to get information indicative of corruption is from data on pharmaceutical prices. The conceptual basis for this approach is that if kickbacks are given, they would result in higher procurement prices. Such an approach has been pioneered by Di Tella and Savedoff (2001) in their study of Latin American hospitals. We collected data on the prices of the following five commonly used medicines for hospital in-patients: ciproflaxin, methoclopramide, diazepam, pentoxyfillene, and amikacin. Hospitals were asked to report how much they paid per unit, how many units they purchased and asked

to specify the units they were referring to (e.g. tablets or packages) for the year 2003. In fact, there were large variations in the prices that hospitals reportedly paid for the same units, sometimes by a factor of 10 or more. However, some patterns in the data suggested that these variations may be due to hospitals misreporting the units (e.g., some hospitals reported paying approximately 0.40 levs for a tablet of ciproflaxin and most others reported paying 4.00. Correspondingly, two hospitals reported paying around 0.40 lev for a package of ciproflaxin and most others reported paying around 4 levs per package. Thus, rather than paying 10 times as much for ciproflaxin as other hospitals, some hospitals may simply be misreporting the units). For this reason, it was difficult to infer the presence of improprieties by an examination of the price data. A related examination was the study of whether hospitals are buying too much or too little medicine for their size and specialty. Again this effort was stymied by the likely misreporting of units.

We did an external cross-check, using audit reports on the hospitals. We found this to be a challenging task. IHHII senior staff had to travel across Bulgaria to get physical access to the voluminous audit reports. The reports themselves lacked summaries and senior project members at IHHII had to read the audit reports to provide us with assessments of the presence of improprieties. Hence, we received an analysis of the audits for only 25 of the 148 hospitals in our study. IHHII graded the audit reports on a 1-5 scale where 1 corresponded to no improprieties and 5 to very serious improprieties. These data were correlated with the aggregate corruption measure derived from the doctors and nurses surveys. This gives some credibility, to the corruption variable constructed using the doctors and nurses data. Had we been able to collect data on the audit reports of all hospitals we would have been able to conduct a more reliable consistency check. In addition we may have been able to evaluate which of doctors and nurses are providing us with better data and given the more reliable source a greater weight. Indeed, we would have been able to use the audit data itself as an outcome variable.

The methodological lesson from the external cross checks is that substantial resources need to be allocated to external cross checks. Had we been able to solve the units problem and been able to conduct the price and quantity analysis on all hospitals, or had the time to conduct audit reports of all hospitals, we would probably have been able to do a better study and produce clearer results.

Construction of Integrity (TAPEE) Indices

The next step in the planned process was the construction of indices of Transparency, Accountability, Prevention, Enforcement, and Education. These indices were created by combining information on lots of specific questions. Our own understanding of the meaning of each of these terms had evolved over the project, as had our understanding of the extent to which they overlapped. A team of three members of IRIS staff discussed the questions on the evaluation committee questionnaire and assigned questions to the 5 TAPEE categories. The answers to each of these questions were then reordered so that higher numbers corresponded to integrity, and the variables were averaged to produce indices for each of the TAPEE categories.

The TAPEE variables did not show much internal coherence. If each of the (say) Transparency variables was reflecting some deeper notion of transparency, we would expect the different variables to be correlated. Hospitals doing better on some aspect of transparency would also (on average) do better on others. However, there was not much evidence of this. Thus while there was agreement on specific items in the (say) transparency index, transparency itself, as a concept, appears to be poorly measured. Similar statements can be made about other aspects of integrity. This may have been one of the reasons why the statistical results linking corruption to integrity were weak.

Regression Analysis of Corruption and TAPEE

In the final stage we ran correlations and regressions between the TAPEE variables and corruption. We only used the measure of corruption from doctors and nurses because we wanted a representative data set.

In broad terms, we did not find any significant results. Corruption is not correlated with any of the TAPEE variables, either in piecewise correlations or in a regression in which we include all the TAPEE variables. Nor does it appear that this problem is due to the way we have aggregated the integrity variables. The corruption variable is individually correlated with only 2/35 of the individual integrity questions (as many as we would expect just by chance). The first of these questions is whether the evaluation committee member attended any training in ethics; the second question is on the severity of punishment if caught taking bribes.

There are several reasons why we may not have found a significant relationship between integrity and corruption. It is possible that we simply didn't collect data on the right aspects of integrity, but this seems unlikely because we did ask a variety of respondents and health experts during the pre-test period whether we were missing out on important aspects of integrity and were told that we had captured the most important aspects. The data may be contaminated by persistent misreports on corruption and transparency despite our efforts to weed out the reticent respondents. Or the doctors and nurses, from whose answers we created the corruption index, may be too uninformed about the procurement process. Thus the absence of results does not imply that the TAPEE variables are unimportant. But we do need to acknowledge we were not able to find a significant empirical relationship between corruption and TAPEE variables in Bulgarian hospitals.

Methodological lessons

The methodological lessons we learnt from the survey and analysis are

It is vital to assess the reticence of respondents. A significant proportion of respondents appear to be reticent. These should be identified and deleted from the sample with the

method used in this study, or the related method used in the Romania study (Azfar and Murrell 2005).

It is important to collect data from multiple respondents at each micro-organization for two reasons. First, so that reticent respondents can be dropped without losing all information on the micro-organization. And second, so that internal cross-checks can be conducted.

It is important to conduct external cross checks and to harness to significant resources that are needed to collect non-survey data on the entire sample.

Conclusion

In bringing together our findings from the two phases of the study, we aimed to address two points: (i) the causal nexus of institutional integrity and corruption; and (ii) the link between pressures and practices at the level of central listing, on one hand, and results at the level of hospital procurements on the other hand. This analysis would then produce an encompassing view of governance quality in Bulgaria's pharmaceutical system.

On the first point, we used qualitative methods at the central level, and regressions of TAPEE and corruption indices at the hospital level. No statistical correlation was found, which suggests that this linkage may not be as strong as USAID initially believed. On the other hand, our analysis suggests that integrity factors play a role among several major influences that lead to (or away from) corruption.

The second point could not be rigorously tested, in large part due to the inadequate quality of pricing data at the hospital level. This prevented us from checking correlations between pricing anomalies in selection and procurement. Also, a comparison of anomalies in drug choice at the central level with those at the hospital level – which might have yielded information on rational drug use – were not feasible due to time and resource constraints. Analytically, corrupt practices at the central and hospital level should be linked in the same way that professional and marketing practice at the two levels are linked. Testing this in future might well be useful, although it would require both substantial resources and increased transparency on the part of NHIF and the hospitals.

To conclude, we tried in this study to couple the assessment of integrity and corruption at the “micro-institution” level of hospitals with an assessment of integrity and (grand) corruption at the central level – supplementing the latter with analysis of “drivers” and outcomes. This was a complex inquiry that produced hard statistical evidence of some factors and practices, and less rigorous probabilistic evidence of higher-level practices and linkages. It also produced a wealth of policy-relevant findings and recommendations. In terms of methodology, we would suggest that this form of inquiry is promising enough to merit further refinement and testing.

References

- Azfar, Omar and Peter Murrell (2005) Identifying Reticent Respondents: Assessing the Quality of Data on Corruption and Values, mimeo, IRIS, University of Maryland, College Park .
- Cohen, Jillian Clare, James Cercone and Roman Macaya. (2002), “Improving Transparency in Pharmaceutical Systems: Strengthening Critical Decision Points Against Corruption.” World Bank Latin American Caribbean Region, Human Development Network, Washington, D.C.: mimeo.
- Management Sciences for Health with the World Health Organization (1997), *Managing Drug Supply: The Selection, Procurement, Distribution, and Use of Pharmaceuticals*. Second Edition, Kumarian Press; West Hartford.
- World Health Organization. (2003), “Annual Report 2002. Essential Drugs and Medicine Policy: Supporting countries to close the access gap.” www.who.int/medicines/library/edm_general/annual_rep/ar2002.pdf. Retrieved March 12, 2004.

Annex 1. Summary Protocol, Drug Selection Component

Question	“Best practice” benchmark ³
Transparency	
1. Are selection guidelines and inclusion/exclusion criteria published and available? Are they clear?	Explicit criteria must be defined and published. Final selection criteria should be based on discussions and acceptance by key prescribers. (See WHO criteria for the selection of essential drugs).
2. Is the following information about committees and officials making selection decisions published and available: their names, basis of appointment, responsibilities?	Names of selection committee members, their qualifications, and their terms of reference should be public information and listed in the formulary manual and on a government website. The method of appointment should also be clearly stated and publicly available. An organigram which is also publicly available should document each member’s background and responsibilities.
3. How do stakeholders learn about decisions?	Announcement of decisions at public meetings, and an information system that disseminates drug selection criteria and rationales helps to ensure integrity and that, if improprieties take place, they are detectable.
4. Are the drug selection meetings open to the public? Announced in advance? In fact attended and covered by the media?	Public scrutiny of drug selection meetings contributes to transparency and limits unethical practices. Media coverage helps ensure transparency and public knowledge of the processes and decisions.
5. Are selection processes documented, and are the records publicly available?	Minutes of selection committee meetings should be archived and available to the public
Accountability	
1. Are drug selection criteria evidence-based? Are the criteria respected in practice?	The government should have clear guidelines that specify what criteria are being applied for drugs on any public formulary. A transparent methodology that determines the drugs’ necessity for the health needs of the population and cost-effectiveness should be uniformly applied. Drug selection must be matched with the pattern of prevalent diseases in country. Government should maintain an information system that monitors drugs once they are in the market.
2. Are choices in the selection process explained (e.g. inclusion, exclusion, deletion)? Are these explanations publicly available?	Formulary drugs should be listed by generic name. Where possible, generic drugs should be used. The inclusion of a new drug should be based on studies that confirm that the drug is necessary for the health needs of the population and on cost-effectiveness. This is particularly relevant for drugs that are not essential drugs. Deletion of drugs from the national drug formulary should be based on sound evidence that they are inappropriate or not cost-effective for the health needs of the population.
3. What forms of official oversight of this process exist, in principle and in practice? How stringent are they?	Selections are best made by an independent commission of professionals that is subject to oversight by some combination of the public, the health professions, the courts (administrative law review), by supreme audit agency, and parliament.

³ Source: Cohen, Cercone and Macaya (2002), MSH & WHO (1997).

4. In what ways can the public provide input to these processes, e.g. applications, appeals, review and comment on proposed rules?	Open and formal consultations with the public should be institutionalized to ensure that all stakeholder views are taken into account in the drug selection process and that no one group has undue influence. There should be a formalized and regular appeal process for applicants who have their drug submissions rejected, to ensure that standards of drug selection are transparent and fair.
Prevention	
1. How and by whom are drug selection officials appointed? How long is their tenure?	The drug formulary committee could be the national drug committee or a smaller subcommittee of it. The appointment process should be public and subject to inputs from a number of persons. The committee membership should be rotating or limited in time to reduce likelihood for systematic bias in the decision making process and to limit individuals power and influence in decision making.
2. Do the committees and officials who make selections have the appropriate mix of skills? Are they neutral, or do they represent a balance of stakeholder interests?	The committee should be formally established and composed of professionals with the requisite technical skills, and meet on a regular basis. It should ideally include a clinical pharmacist or pharmacologist, a physician, economist and medical specialists who can prepare and/or review drugs.
3. What other occupations and activities are selection officials involved in – including active medical practice? Do the rules require the declaration, or at least the avoidance, of possible conflicts-of-interest? Are there limits to officials' contacts with drug companies?	Committee members should disclose all other involvement that may be perceived as conflict of interest. If overlapping responsibilities suggest conflict of interest, the committee member should be compelled to either give up a particular role or resign. Committee members should not have active medical practices, to avoid conflict of interest. Committee members should declare any personal conflicts of interest in writing. These statements should be publicly available.
4. Are drug-selection procedures conducted regularly, or are there delays between sessions?	Drug selection committee meetings should take place on a set schedule. This will help promote reasonable timelines for decision making and more transparency. There should be minimal delays for market authorization and selection decisions if sufficient information is presented to the government institution.
5. What methods are used to make selections, e.g. unanimous decision, majority vote, choice by individual official? Are decisions vulnerable to political influence – and how is this addressed?	Decision making should be democratic, transparent and subject to formalized voting procedures that rely on majority for outcomes. There are four major methods for quantifying drug needs: consumption (based on historical data), morbidity based, adjusted consumption, and service-level projection. Ideally, a combination of these will be applied to obtain the most accurate drugs for the health needs of the population.
6. Can interested firms influence the selection process? What methods do they use – e.g. policy arguments, education and promotion, meetings with relevant officials, favors?	There should be clear laws, code of conduct, and regulations governing industry marketing practices. Officials who are involved in drug selection decisions should be barred from meeting with drug company representatives to avoid any potential conflict of interest of influence on decision-making. The government should have a law that explicitly prevents public officials who are members of the drug selection committee from accepting gifts in cash or kind from pharmaceutical companies.
Enforcement	
1. Are the rules on official appointments and terms of reference respected in practice?	Clear, public, and well-enforced appointment rules and terms of reference for each drug selection committee should be in place.

<p>2. What sanctions are there for breach of the rules on conflict-of-interest? Bribery and other forms of corruption?</p>	<p>Well-defined sanctions should be applied if a committee member engages in inappropriate (unethical) conduct. By enforcing sanctions appropriately and effectively, this will also serve as a deterrent to any future misguided actions. In most countries, bribery legislation is included in the penal code or in special corruption legislation.</p>
<p>3. Are there mechanisms in place to detect improper relationships – e.g. selection officials with undisclosed economic interests in the pharmaceutical sector? Are these effective in practice, or are such relationships accepted?</p>	<p>Any member on a drug selection committee should have no connections (formal or informal) to a pharmaceutical company. Committee members and external experts working with them should disclose all other involvement that may potentially create a conflict of interest. If overlapping responsibilities suggest conflict of interest, the committee member/expert should be compelled to either give up a particular role or resign. Public officials should have the duty, and the information necessary, identify if companies bidding for the same tender have any corporate relationships.</p>
<p><i>Education</i></p>	
<p>1. Do selection committees or officials inform, educate, or solicit input from stakeholders?</p>	<p>Drug selection committee members should regularly organize public education campaigns and consultations to ensure fair input on decision making and procedures.</p>
<p>2. How are these and other relevant officials trained in ethics and integrity rules? How stringent are these rules, in principle and in practice?</p>	<p>All drug selection officials should be trained regularly on ethical guidelines, standards of practice and consequences for any breaches.</p>

Self administered version

You have two envelopes in your survey packet, Envelope 1 and Envelope 2.

Each of these two envelopes contains two questions. One is innocuous and the other is about corruption. One of those two questions is underlined. You are supposed to answer the underlined question. More than 10% of the respondents have the innocuous question underlined.

First you should open Envelope 1 and answer the underlined question. Do not show the sheet to anyone! Do not tell anyone which question is underlined! Just see which one it is and give an answer to it. The methodology of the research requires you should not comment these questions with other people.

Please, just give an answer to the underlined question, then put the sheet back into the envelope and keep it to throw it away later.

60. Open Envelope 1. Answer for Envelope 1.

Answers:

1. Yes

2. No

Now you should proceed in the same manner with Envelope 2.

61. Open Envelope 2. Answer for Envelope 2.

Answers:

1. Yes

2. No

Envelope Method

printed on the envelope:

Fact Marketing Agency
Envelope 1.

In 90%

Have you ever stolen money or medicine?

Is $3+3 = 6$?

In 10%

Have you ever stolen money or medicine?

Is $3+3 = 6$?

printed on the envelope:

Fact Marketing Agency
Envelope 2.

In 100%

Did you ever accept a bribe?

Is $2+2 = 4$?