Safeguards verification – its credibility and the diversion hypothesis

by H. Gruemm*

The main purpose of IAEA safeguards as a verification system is to provide *assurance* that States comply with their commitments in relation to their peaceful use of nuclear energy. Secondly, IAEA safeguards act as a warning system, and may thus constitute a *deterrent* to contemplated diversion by creating the risk of early detection. To achieve their purpose, IAEA safeguards must be *credible:* they must be not only effective, but perceived to be effective. This puts the focus on safeguards effectiveness. It is however very difficult to define and, in particular, to quantify this basic requirement of the verification system.

Taking into account the fact that States conclude safeguards agreements by their own free decision, and considering other circumstances, it seems reasonable to assume that only a few States, if any, might ever contemplate diversion. This is one of the consequences of world-wide adherence to the non-proliferation idea supported by IAEA safeguards. Paradoxically, effective safeguards contribute to the difficulty of measuring safeguards effectiveness by the most simple indicator, namely the percentage of diversion acts or related events detected during a given period.

Under these circumstances the only way of making the effectiveness of safeguards visible rests on the periodic critical review of safeguards activities. This review should ideally demonstrate that IAEA verification activities are so thorough that diversion of nuclear material or misuse of facilities would be detected with high probability should it occur. This means that the IAEA, in developing an effective verification methodology, has to assume as a general *working hypothesis* that diversion cannot be excluded and that consequently a diversion risk of low but non-zero probability exists in all cases of safeguards implementation[1].

The design of IAEA verification activities as document audits, measurements, and observations, at nuclear facilities is based on this working hypothesis. If these activities are well planned and executed and lead to the result that the diversion hypothesis cannot be proven, one can conclude with a high level of confidence that, in fact, no diversion has occurred. Assurance thus results from conclusions arrived at by thorough IAEA verification activities[2]. Conceptually IAEA verification can thus be regarded as the testing of diversion hypotheses.

The diversion hypothesis obviously plays an important rôle in designing and organizing effective and credible verification activities. Therefore, any diversion analysis has to consider a wide range of potential "diversion strategies" and possible concealment methods for various types of nuclear material and facilities. Such an analysis includes consideration of the technical characteristics of the nuclear facility, of the type and location of nuclear material, and of possible diversion paths as well as possible diversion rates.

It would not be very realistic to consider scenarios in which an inspector detects the specific act of diversion: that is, catches an operator red-handed removing material from authorized uses. It is therefore the purpose of the diversion analysis to identify *anomalies*[3], that is to say "observables", that might be indicative of acts of diversion. Safeguards approaches are then designed to ensure that verification activities focus on anomalies and provide an adequate detection probability.

There may be different causes for the occurrence of anomalies. Normally they result from entirely *innocent causes* and many such anomalies are found each year.

---

* Mr Gruemm is the former Deputy Director General of the IAEA Department of Safeguards.

[1] This hypothesis should not be understood – and in general is not understood – as an expression of distrust directed against States in general or any State in particular. Any misunderstanding might be dispelled by comparing the diversion hypothesis with the philosophy of airport control. In order to be effective, airport control has to assume *a priori* and without any suspicion against a particular passenger that each handbag might contain prohibited goods.

[2] Such a conceptual approach to verification in the context of international safeguards is also contained implicitly in IAEA document INFCIRC/153/(Corr.) providing in para.19 that if the Board of Governors, upon examination of relevant information reported to it by the Director General, finds that the IAEA is unable to verify that there has been no diversion of nuclear material, it may make reports to Member States, the Security Council, and the General Assembly and may take other measures against the State.

[3] Examples are inconsistencies in documents, inaccessability of nuclear facilities, IAEA seals tampered with, etc.

They have various causes such as:

- printing or calculation errors in records or reports;
- incomplete records or reports;
- measurement errors;
- inadvertent interference with IAEA seals or instruments;
- failure of IAEA equipment; or
- errors made by the inspector.

Anomalies also could be the consequence of *diversion* or intended diversion resulting from, for example:

- the unreported removal from or introduction into a nuclear facility of nuclear material, including the case where the material might come from or flow into a nuclear facility not subject to safeguards;
- unreported change of composition of nuclear material within the facility (e.g. production of plutonium from fertile material, or enrichment of nuclear material above the declared level); or
- proscribed uses of nuclear material within the facility[4].

For this reason the IAEA in its verification work uses follow-up activities which are intended to resolve each anomaly found and to ensure that no true alarm sign is ignored or false alarm raised. If all anomalies found have been satisfactorily explained, the IAEA can state as an objective fact that during the reporting period no anomaly was detected which would indicate diversion. The thoroughness of the verification methods applied permits one then to conclude with a high level of confidence that also in reality no such anomaly existed and that therefore no diversion occurred.

One of the elements of the hypothesis to be tested by verification includes the possibility that a divertor might try to *conceal* anomalies created by diversion. Concealment actions to be taken into account in designing safeguards approaches may include, inter alia:

- the falsification of records and reports by understating receipts or production of nuclear material, or by overstating shipments, losses or inventories;
- the replacement of missing material by material of lower safeguards significance or by material, borrowed from other facilities;
- the manipulation of measurements or of their evaluation; or
- interference with containment or with IAEA equipment.

As far as the significant elements of diversion hypotheses are concerned, discussions have shown that some safeguards experts have difficulties with the inclusion of certain concepts in the diversion hypothesis, in particular the existence of "clandestine facilities" in a full-scope safeguards situation and of "unreported nuclear material".

Under INFCIRC/66/Rev.2 safeguards agreements, unsafeguarded facilities may exist in a State. Their existence is taken into account in developing the safeguards approach for the facilities under safeguards.

NPT safeguards agreements according to INFCIRC/153/(Corr.) give the IAEA the ". . . right and obligation to ensure that safeguards will be applied, in accordance with the terms of the agreement, on all source or special fissionable material in all peaceful nuclear activities within the territory of the State . . .". The "terms of the agreement" include verification procedures by the IAEA to ensure compliance with the basic commitment by the State, namely *not to divert nuclear material.* However, no specific procedures are foreseen for verification by the IAEA of the second commitment, namely the *reporting of all nuclear material subject to safeguards,* in particular for verification of the completeness of the initial inventory report. The reasons for this are understandable: a kind of international police organization with inspectors roaming around in sovereign States in the search of possible clandestine nuclear facilities or material is universally unacceptable and has not been suggested by anyone.

As a consequence of its inability to verify the completeness of reports by States, the IAEA cannot exclude from realistic scenarios the hypothesis that *unsafeguarded facilities* which are connected by a flow of nuclear material with safeguarded facilities might exist also in an NPT situation. For instance, at least an assembly workshop is required for the manufacture of a nuclear explosive and it can obviously not be expected that this would be submitted to safeguards. If the possibility of the existence of such a facility can be excluded *a priori,* the manufacture of a nuclear explosive would be physically impossible. As a consequence the reason for verifying declared material would become meaningless.

Also, the existence of a hot cell complex not containing nuclear material - and therefore not under safeguards - that might be used for the reprocessing of diverted spent fuel cannot be excluded in a credible diversion hypothesis. Nor can the diversion and stockpiling of spent fuel for later prohibited uses be excluded *a priori.* Unless assumptions of this kind were made, it could be argued that no or little purpose would be served by implementing safeguards in a country which had only power reactors, because separation of plutonium from spent fuel or enrichment of uranium – and thus manufacture of nuclear explosives - would then be physically impossible.

It follows that in analysing possible diversion scenarios the IAEA has to assume as a *working hypothesis* that diversion paths might lead from facilities containing safeguarded material to unsafeguarded facilities and *vice versa.* Only then can assurance be provided that all anomalies at a safeguarded facility which might indicate diversion have been properly identified, and that the non-detection of any significant anomaly means that, in fact, no diversion occurred. Verification measures to resolve all anomalies disclosed at a facility whatever their

[4] Diversion of nuclear material from peaceful uses does not necessarily mean removal of the material from a facility.

cause may be are foreseen in the relevant facility attachment and are covered by the terms of the safeguards agreement. The application of such measures which have been worked out with that purpose in mind is part of the obligation of the IAEA to apply safeguards to *all* peaceful nuclear activities in a country.

Certain misunderstandings have also arisen from the assumption in IAEA diversion analysis that the existence of *unreported material* cannot, *a priori,* be excluded. However, the most simple diversion scenario consists of the unreported clandestine removal of spent fuel from a power plant and its introduction into a reprocessing plant in the hope that the separation of plutonium and its subsequent removal will not be detected.

Nuclear material on a diversion path remains legally subject to safeguards, even if it is not reported by the diverter. Under the agreements IAEA inspectors have the right to verify that no such material is entering or leaving the nuclear facilities and to make the necessary enquiries should doubts arise. The identification of anomalies which might indicate the introduction of unreported material into a facility[5] is, therefore, also one of the essential tasks in developing safeguards approaches. Similar considerations would apply with respect to the possibility that unreported fertile material might be introduced in a reactor for purposes of plutonium production and the subsequent removal of the irradiated material and separation of plutonium. Such action would moreover violate the States' commitment to report to the IAEA any change in the design of a facility.

As in all other cases the IAEA applied this concept in the case of the OSIRAK research reactor in Iraq: the possibility of the undeclared production of plutonium and of the existence of clandestine fuel production and reprocessing capabilities had to be assumed in preparing the safeguards regime to be applied after start-up of the reactor. This regime foresaw frequent inspections and the installation of automatic cameras to detect the possible clandestine introduction of fertile material and removal of irradiated material. It is not hard to imagine the loss of IAEA safeguards credibility which would have followed the attack on the OSIRAK reactor if the IAEA had heeded some of its critics and not made the above assumptions.

Referring to the question of quantification of effectiveness as an essential element of safeguards credibility, there is no methodology yet available to characterize safeguards effectiveness in a quantitative way. Other indicators of effectiveness must therefore be analysed. Safeguards effectiveness is obviously connected directly with two factors:

- *the scope of achievement* resulting from safeguards implementation: this can be defined as the percentage of nuclear material/facilities under safeguards for which the inspection goals have been fully attained, and
- the *level of assurance* attained, which depends, among other things, on the overall probability of detection of a diversion, if one had occurred.

Thus, both scope of implementation and level of assurance are principal factors in assessing the effectiveness of safeguards verification. It is not too difficult to determine the scope of achievement. The proportion of nuclear facilities where the inspection goals have been attained has increased considerably over the last years. However, the level of assurance cannot be derived directly from the degree of inspection goal attainment, because the level of assurance is related to the detection probability, which is only implicitly contained in the criteria used for the evaluation of goal attainment. It should be noted that the evaluation of safeguards effectiveness has become more stringent over the years as more manpower and equipment have become available and more systematic evaluation methods have been introduced. As a consequence, the completeness and intensity of coverage of plausible diversion paths and of concealment methods, as well as the overall detection probability, have increased considerably.

In assessing the effectiveness and thereby the credibility of IAEA safeguards it should not be forgotten that the reports and conclusions of the IAEA are not the only source of information available to Member States. They may have their own national means for detecting unsafeguarded nuclear activities; they may take into account the internal and external situation of States and assess their political intentions and their technological capabilities.

[5] Examples of such anomalies are: a substantial positive amount of nuclear material unaccounted for (MUF), or a discrepancy between receipts of fuel assemblies reported by a reprocessing plant and the number of crane movements established by IAEA surveillance.