February 2005

AVIATION SECURITY

Measures for Testing the Impact of Using Commercial Data for the Secure Flight Program
February 23, 2005

Congressional Committees:

The Transportation Security Administration (TSA) is developing a new passenger prescreening program, known as Secure Flight. Under the Secure Flight program, TSA plans to take over, from commercial airlines, the responsibility for comparing identifying information of domestic airline passengers against information on known or suspected terrorists. TSA is also considering using commercial data as part of Secure Flight if the data are shown, through testing, to improve the results of these comparisons.¹ In the 2005 Homeland Security Appropriations Act (Public Law 108-334, Section 522(d)), Congress mandated that, prior to testing the use of commercial data for Secure Flight, TSA develop measures to assess the impacts of using commercial data on aviation security, and that GAO review the measures. In response to that mandate, we reviewed TSA’s measures for commercial data testing and briefed congressional staff on January 11, 2005, on our findings. This report documents the results of our review, which we presented in that briefing.

Currently, commercial airlines are responsible for the prescreening of passengers using terrorist watch lists provided by TSA, known as the no-fly and selectee lists.² However, as noted by the National Commission on Terrorist Attacks Upon the United States (9/11 Commission), the watch lists used by the airlines do not include all terrorists or terrorism suspects because of concerns about sharing intelligence information with private firms and foreign countries.³ TSA expects that Secure Flight will improve passenger prescreening as compared with the current airline-operated process. For example, Secure Flight will utilize an expanded terrorist watch list that includes information not currently provided to air carriers.

¹Commercial data are maintained by private companies and can include personally identifiable information that either identifies an individual or is directly attributed to an individual, such as name, address, and phone number.

²To conduct passenger prescreening, airlines also compare passenger data against the Computer-Assisted Passenger Prescreening System (CAPPs I) rules, which are behavioral characteristics associated with the way an airline ticket is purchased. The CAPPs I rules are intended to identify individuals who should receive additional security scrutiny.

for passenger prescreening. TSA also expects that by automating the prescreening process and applying consistent procedures for comparing passenger data against the expanded terrorist watch list, Secure Flight will reduce the number of false positive matches against the terrorist watch list as compared with the current process.

In preparing to take over passenger prescreening from domestic air carriers, TSA has begun initial Secure Flight testing to determine the ability of Secure Flight to effectively compare passenger-provided information contained in air carrier reservation systems against the expanded watch list in order to identify individuals known or reasonably suspected to be engaged in terrorism. TSA expects that results from these tests will be available in February 2005. In addition, TSA plans to conduct a concept test to determine if the use of commercial data can improve the matching of passenger-provided information against the expanded watch list by identifying individuals who were incorrectly identified as being on a terrorist watch list (referred to as false positives) or who attempted to avoid detection by disguising their identity (referred to as false negatives). The commercial data concept test is also intended to determine if the accuracy of passenger-provided data contained in passenger records can be verified using commercial data. In January 2005, TSA issued a request for proposals in order to obtain a contractor to conduct commercial data concept testing. TSA expects to award the contract in late February 2005.

To determine the effectiveness of using commercial data, TSA developed initial measures for commercial data concept testing, such as the overall percentage of passenger-provided records from which identity can be verified using commercial data, and plans to refine the measures throughout the testing process. TSA expects to obtain the results of commercial data concept testing in April 2005. On the basis of these test results, the Department of Homeland Security (DHS) and TSA plan to

---

4These reservation systems contain detailed information about an individual's travel on a particular flight, including information provided by the passenger when making a flight reservation. Such information can include (1) passenger name, (2) reservation date, (3) travel agency or agent, (4) travel itinerary information, (5) form of payment, (6) flight number, and (7) seating location.

5The purpose of the concept test is limited to identifying the utility of using commercial data in improving the effectiveness of comparing passenger information against the terrorist watch list in a test environment.

6To obtain data for Secure Flight testing, TSA issued an order in November 2004 requiring domestic airlines to provide passenger records for the month of June 2004.
make policy decisions regarding the use of commercial data as part of the overall Secure Flight program. TSA also plans to subsequently test additional functionality and the operations of Secure Flight before implementation, regardless of whether TSA incorporates the use of commercial data as part of Secure Flight.

To determine if the measures developed by TSA for commercial data testing are designed to identify impacts on aviation security, we reviewed and analyzed TSA’s draft statement of work for commercial data concept testing, which includes the initial measures developed by TSA. Since the purpose of our review was to determine whether the measures identify impacts on aviation security, we assessed the measures against performance measurement criteria developed by GAO based on best practices. On the basis of our knowledge of the Secure Flight program and GAO performance measurement criteria, we determined whether TSA’s measures are designed to reflect relevant impacts on aviation security and are consistent with attributes of successful performance measures. We also interviewed TSA officials responsible for Secure Flight development and oversight. The briefing slides, contained in appendix I, include the specific attributes that we used as criteria for evaluating TSA’s measures, detailed information on our scope and methodology, and the results of our review of TSA’s measures for commercial data testing. Appendix I also includes a list of TSA’s initial measures for commercial data testing. We conducted our work in accordance with generally accepted government auditing standards from December 2004 to February 2005. GAO is also continuing to review TSA’s measures for commercial data testing based on a follow-on congressional request.

In January 2005, we briefed your offices on the results of our review of TSA’s measures for commercial data concept testing. In summary, we made the following key points in our briefing:

- TSA developed a concept test to determine the utility of using commercial data for Secure Flight as a first step in determining its impact on aviation security. The results of this test are intended to provide TSA the basis for

---

7 Performance measurement is used to provide information on the achievement of program accomplishments, particularly progress toward meeting preestablished goals or targets, and the impacts of those accomplishments.

8 TSA’s final statement of work for commercial data testing, issued subsequent to our briefing, includes a revised set of measures for the use of commercial data. We will assess these revised measures as part of our follow-on review of TSA’s commercial data test.
refining performance measures identifying impacts on aviation security prior to subsequent testing, should DHS and TSA decide to pursue the use of commercial data.

- TSA developed initial measures for commercial data concept testing that are intended to provide information related to impacts on aviation security, including improvements in false positive and false negative rates. TSA, in coordination with the contractor, plans to refine these measures during concept testing—to include the establishment of performance targets—and prior to operationally testing the system, should DHS and TSA decide to pursue the use of commercial data.

- TSA measures developed to date for commercial data testing do not, and were not designed to, provide information on overall Secure Flight system operations (i.e., system response time, connectivity with air carriers, security, and privacy) or identify impacts of using commercial data on aviation security in an operational environment. Accordingly, the measures do not generally reflect attributes of successful performance measures for this purpose.

- Additional work reviewing TSA’s refined measures, should DHS and TSA decide to pursue the use of commercial data for Secure Flight, would be needed to determine if the measures are designed to identify relevant impacts on aviation security, and reflect attributes of successful performance measures for that purpose.

We provided a draft of this report to DHS for its review and comment. In commenting on the draft report, DHS generally agreed with our findings. DHS's written comments are presented in appendix II. TSA also provided technical comments which we have incorporated into this report where appropriate.

We are sending copies of this report to the Secretary of the Department of Homeland Security and the Administrator of the Transportation Security Administration. We will also make copies available to others upon request. In addition, the report will be available at no charge on GAO’s Web site at http://www.gao.gov.
If you or your staff have any questions about this report, please contact me at (202) 512-3404 (berrickc@gao.gov) or Christine Fossett, Assistant Director, at (202) 512-2956 (fossettc@gao.gov). Other key contributors to this report were R. Denton Herring, Adam Hoffman, David Hooper, Tom Lombardi, and David Plocher.

Sincerely yours,

Cathleen A. Berrick
Director, Homeland Security
and Justice Issues
List of Congressional Committees

The Honorable Thad Cochran
Chairman
The Honorable Robert C. Byrd
Ranking Minority Member
Committee on Appropriations
United States Senate

The Honorable Ted Stevens
Chairman
The Honorable Daniel K. Inouye
Ranking Minority Member
Committee on Commerce, Science, and Transportation
United States Senate

The Honorable Jerry Lewis
Chairman
The Honorable David R. Obey
Ranking Minority Member
Committee on Appropriations
House of Representatives

The Honorable Don Young
Chairman
The Honorable James L. Oberstar
Ranking Minority Member
Committee on Transportation and Infrastructure
House of Representatives

The Honorable Tom Davis
Chairman
Committee on Government Reform
House of Representatives

The Honorable Adam H. Putnam
House of Representatives
Appendix I: Briefing Slides

Review of TSA Measures for Secure Flight Commercial Data Testing

Briefing for the Majority and Minority Staff of the Cognizant Senate and House Authorization, Appropriations, and Oversight Committees

January 11, 2005
Briefing Outline

• Background
• Objective, Scope, and Methodology
• Summary
• TSA Secure Flight Testing Approach
• Concept Testing for Use of Commercial Data
• Attributes of Successful Performance Measures
• TSA Measures for Commercial Data Testing
• List of TSA Measures for Commercial Data Testing from Draft Statement of Work
Appendix I: Briefing Slides

Background

The Transportation Security Administration (TSA) is developing a new passenger prescreening program, known as Secure Flight. Under this program:

- TSA will assume responsibility for checking airline passengers' names against no-fly and selectee lists and the Computer-Assisted Passenger Prescreening System (CAPPS I) rules, a task that is currently performed by air carriers.\(^1\) TSA will also check passenger names against an expanded terrorist watch list.

- TSA is also considering the use of commercial data (e.g., personally identifiable information that either identifies an individual or is directly attributed to an individual, such as name, address, and phone number) if it is shown, through testing, to be effective in mitigating false positives, identifying false negatives, or verifying passenger identification.\(^2\)

---

\(^1\) CAPPS I rules are behavioral characteristics used to select passengers who require additional security scrutiny at airport security checkpoints.

\(^2\) A false positive is an individual who was misidentified as a positive match when matching passenger data against a terrorist watch list. A false negative is an individual on a terrorist watch list who avoids detection when passenger data are matched against a terrorist watch list by disguising his or her true identity.
Background (cont’d.)

TSA has two testing efforts for the Secure Flight program:

- watch list/CAPPS I testing—a test to match historical passenger data against an expanded government watch list and CAPPS I rules.

- commercial data concept testing—a test of a methodology to determine if the use of commercial data can improve on the results of watch list/CAPPS I testing.

At the conclusion of both tests, TSA plans to conduct additional Secure Flight system testing.
Objective, Scope, and Methodology

Objective

- Evaluate TSA measures for Secure Flight testing using commercial data to determine if they are designed to identify impacts on aviation security as required by Public Law 108-334, Section 522 (d).

Scope and Methodology

- Reviewed TSA’s draft statement of work, which includes the initial measures developed by TSA, and the draft request for proposals for commercial data concept testing, and interviewed TSA officials responsible for Secure Flight program development and oversight.

- Determined, on the basis of our knowledge of the Secure Flight program and performance measurement criteria GAO developed based on best practices (shown on slide 13), whether the measures are designed to reflect relevant impacts and whether they were prepared in accordance with attributes of successful performance measures.
Scope and Methodology (cont’d.)

- Utilized GAO personnel expert in performance measurement, contracting, and information systems development and testing issues.

- Conducted our work from December 2004 through January 2005 in accordance with generally accepted government auditing standards.
Summary

- TSA developed a concept test to determine the utility of using commercial data for Secure Flight as a first step in determining its impact on aviation security. The results of this test are intended to provide TSA with the basis for refining performance measures that identify impacts on aviation security prior to subsequent testing, should the Department of Homeland Security (DHS) and TSA decide to pursue the use of commercial data.

- TSA developed initial measures for commercial data concept testing that are intended to provide information related to impacts on aviation security, including improvements in false positive and false negative rates. TSA, in coordination with the contractor, plans to refine these measures during concept testing—to include the establishment of performance targets—and prior to overall Secure Flight system testing, should DHS and TSA decide to pursue the use of commercial data.

- TSA measures developed to date for commercial data testing do not, and were not designed to, provide information on overall system operations (i.e., system response time, connectivity with air carriers, security, and privacy), or identify impacts of using commercial data on aviation security in an operational environment. Accordingly, the measures do not generally reflect attributes of successful performance measures for these purposes.

- Additional work reviewing TSA’s refined measures, should DHS and TSA decide to pursue the use of commercial data for Secure Flight beyond concept testing, would be needed to determine if the measures are designed to identify relevant impacts on aviation security and reflect attributes of successful performance measures.
TSA Secure Flight Testing Approach

TSA has two testing efforts for the Secure Flight program, as shown in figure 1 on the next slide:

- **Secure Flight test (watch list/CAPPS I)**
  - A test to match historical (June 2004) passenger data collected from air carriers against an expanded government watch list and CAPPS I rules.
  - This test is being conducted by IBM and is expected to be completed in February 2005.

- **Commercial data test**
  - TSA plans to award a contract to test the concept of using commercial data (concept testing). This test has two main objectives: to determine the value of commercial data (1) in mitigating false positives and identifying false negatives and (2) in verifying passenger identification.
  - Based on the outcome of concept testing, DHS and TSA plan to make policy decisions regarding the use of commercial data as part of Secure Flight.

At the conclusion of both tests, TSA plans to award a contract in early March to conduct additional Secure Flight development and expects to commence overall system testing in early June 2005.
Figure 1 — Secure Flight Testing Approach

TSA Secure Flight Testing Approach (cont’d.)

Activities
- Issue request for proposals
- Phase I design methodology
- Phase II test methodology
- Test results (42 working days after award)

Products
- Draft statement of work (initial measures developed by TSA to obtain quantitative data on system concept)
- Level 2, 5 commercial data test plan (more specific measures: 5 working days after award)

Source: GAO.
Concept Testing for Use of Commercial Data

- TSA’s concept testing for the use of commercial data is limited in scope and purpose. TSA officials characterized this as a test of a methodology to determine if the use of commercial data can improve the results of matching passenger-provided information to the terrorist watch list. TSA officials said they plan to use the results of watch list and CAPPS I testing as the baseline for measuring improvements in accuracy achieved during commercial data concept testing.

- TSA officials stated that commercial data concept testing is not designed to provide information on overall system operations (i.e., system response time, connectivity with air carriers, security, and privacy) or identify impacts of using commercial data on aviation security in an operational environment.

- TSA officials stated that if DHS and TSA decide to incorporate commercial data as part of the Secure Flight system based on the results of concept testing, operational and related factors will be tested and measured in conjunction with overall Secure Flight system testing.
Concept Testing for Use of Commercial Data (cont’d.)

- TSA plans a two-phase approach for concept testing using commercial data:
  - Phase 1—the contractor will develop a testing methodology for using commercial data, including developing additional measures to test the effectiveness of commercial data to mitigate false positives, identify false negatives, and verify passenger identification.
  - Phase 2—the contractor will carry out the test methodology using June 2004 passenger data obtained from air carriers and report results to TSA.

- At the conclusion of concept testing, the contractor will provide a separate evaluation of the costs to the government associated with commercial data usage. This evaluation, however, will not include costs to the airline industry.

- The contract period for concept testing is planned to be 6 weeks, at an estimated cost of $500,000.
Attributes of Successful Performance Measures

- Sound performance measurement is important in providing decision makers with information on the achievement of program accomplishments, particularly progress toward meeting preestablished goals or targets, and the impacts of those accomplishments. Performance measurement focuses on whether a program has achieved its objectives, expressed as measurable performance standards.

- GAO has developed, based on prior work on performance measurement, nine attributes of successful performance measures, which are shown in table 1 on the next slide.
Attributes of Successful Performance Measures (cont’d.)

Table 1 — Key Attributes of Successful Performance Measures

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Definitions</th>
<th>Potentially adverse consequences of not meeting attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linkage</td>
<td>Measure is aligned with division and agencywide goals and mission and clearly communicated throughout the organization.</td>
<td>Behaviors and incentives created by measures do not support achieving division or agencywide goals or mission.</td>
</tr>
<tr>
<td>Clarity</td>
<td>Measure is clearly stated and the name and definition are consistent with the methodology used to calculate it.</td>
<td>Data could be confusing and misleading to users.</td>
</tr>
<tr>
<td>Measurable target</td>
<td>Measure has a numerical goal.</td>
<td>Can not tell whether performance is meeting expectations.</td>
</tr>
<tr>
<td>Objectivity</td>
<td>Measure is reasonably free from significant bias or manipulation.</td>
<td>Performance assessments may be systematically over- or understated.</td>
</tr>
<tr>
<td>Reliability</td>
<td>Measure produces the same result under similar conditions.</td>
<td>Reported performance data is inconsistent and adds uncertainty.</td>
</tr>
<tr>
<td>Core program activities</td>
<td>Measures cover the activities that an entity is expected to perform to support the intent of the program.</td>
<td>Not enough information available in core program areas to managers and stakeholders.</td>
</tr>
<tr>
<td>Limited overlap</td>
<td>Measure should provide new information beyond that provided by other measures.</td>
<td>Manager may have to sort through redundant, costly information that does not add value.</td>
</tr>
<tr>
<td>Balance</td>
<td>Balance exists when a suite of measures ensures that an organization’s various priorities are covered.</td>
<td>Lack of balance could create skewed incentives when measures over-emphasize some goals.</td>
</tr>
<tr>
<td>Governmentwide priorities</td>
<td>Each measure should cover a priority such as quality, timeliness, and cost of service.</td>
<td>A program’s overall success is at risk if all priorities are not addressed.</td>
</tr>
</tbody>
</table>

TSA Measures for Commercial Data Testing

- TSA has developed initial measures for commercial data concept testing as part of the draft statement of work. The contractor is to develop additional measures—to include the establishment of performance targets—in consultation with TSA, as part of the test methodology. (See slide pages 16-18 for the list of TSA measures.)

- TSA officials stated that the draft statement of work for concept testing was purposely designed to provide vendors with flexibility to leverage commercial knowledge and expertise in developing proposals and identifying additional measures.

- According to TSA officials, measures for concept testing are intended to provide quantitative data that will be used to make policy decisions regarding the use of commercial data as part of the Secure Flight system.
TSA Measures for Commercial Data Testing (cont’d.)

- TSA’s initial measures for concept testing are intended to provide information related to impacts on aviation security, including improvements in false positive and false negative rates. TSA, in coordination with the contractor, plans to refine these measures during concept testing and prior to Secure Flight system testing, should DHS and TSA decide to pursue the use of commercial data.

- Because of the scope of concept testing, TSA measures developed to date do not identify impacts of using commercial data on aviation security in an operational environment. Accordingly, the measures do not generally reflect attributes of successful performance measures for this purpose.

- Additional work reviewing the refined measures, should DHS and TSA decide to pursue the use of commercial data for the Secure Flight System beyond concept testing, would be needed to determine if they are designed to identify relevant impacts on aviation security and reflect attributes of successful performance measures.
TSA Measures for Commercial Data Testing from Draft Statement of Work

Identity verification measures:

- confidence percentage, or likelihood, of verification, and which characteristics (individually or in combination) best supported verification;
- confidence percentage, or likelihood, of accuracy during verification;
- overall percentage of records on which identity can be verified;
- description of criteria for passenger non-verification, reasons for ambiguity (e.g., missing specific information, incorrect information);
- percentage of records meeting the criteria for non-verification;
- distinguishing characteristics of non-verified identities to ensure fair handling and to ensure profiling is not used;
- commonalities among false positives mitigated and commonalities between false positives not mitigated;
- indicators that prove or disprove identity;
- record data elements that provide the most value for matching to commercial data for identity validation;
- percentage of records that contained the data characteristics that best supported verification; and
- average verification rates using passenger data as compared to industry standards.
TSA Measures for Commercial Data Testing from Draft Statement of Work (cont’d.)

False positive reduction and false negative identification measures:

- percentage of positives matches against the watch list determined to be false positives (mitigated) by using commercial data to amplify passenger-supplied information;
- percentage of positives against the watch list that could neither be confirmed as a positive match or eliminated as false positives through the commercial data amplification process that remained false positives;
- determine what information would have been required to finally clear those positive matches that could not be confirmed or eliminated after amplification;
- record data elements that provide the most value for the mitigation of false positives;
- augment sets of records determined to be false positives with commercial data. Provide the enhanced records to the government to re-run through the Secure Flight Platform to determine whether the additional information was effective in avoiding false positive matches against the watch list; and
- augment sets of records determined to be cleared (no match against the watch list) with commercial data. Provide this information to the government to re-run on the Secure Flight Test Platform to determine if the addition of commercial data resulted in watch list matches not previously identified (false negatives).
Cost evaluation measures:

- total cost to deploy identity verification;
- unit cost per identity verified;
- total cost to deploy false positive watch list match reduction (or false negative match identification);
- unit cost per false positive reduced;
- total cost of identity verification and false positive reduction (or false negative identification); and
- unit cost of identity verification and false positive reduction.
February 11, 2005

Ms. Cathleen A. Berrick
Director, Homeland Security and Justice Issues
U.S. Government Accountability Office
441 G Street, NW
Washington, D.C. 20548

Dear Ms. Berrick:

Thank you for the opportunity to comment on GAO’s draft report entitled, “Measures for Testing the Impact of Using Commercial Data for the Secure Flight Program,” GAO-05-324. TSA generally concurs with this initial report and looks forward to continued cooperation with GAO throughout the Secure Flight Program development process. TSA is confident Secure Flight will meet our goals of improving the security of domestic air travelers and reducing passenger airport screening time while protecting privacy and civil liberties.

In the 2005 Homeland Security Appropriations Act Congress mandated that prior to testing, TSA develop measures to assess the impact of using commercial data on aviation security and that GAO review those measures before TSA tests commercial data (P.L. 108-334). TSA is complying with this directive, and has designed and will continue to refine these measures as it prepares to initiate testing – subject to GAO review and Congressional notification. As GAO described, in addition to the Secure Flight test using watch list information and CAPPS II rules, TSA is planning two phases of concept testing for the use of commercial data. The results of the tests will be used to determine the potential cost, feasibility, and effectiveness of using commercial data for Secure Flight. These tests are scheduled to begin in late February or early March. Following completion of this testing, policy decisions regarding whether and to what extent commercial data will be utilized in Secure Flight will be made.

As the GAO recognized, the measures in this report are designed to evaluate the effectiveness of using commercial data within a test environment. TSA recognizes that additional Congressional review will be required once the results of the commercial data tests become available. However, it is important to emphasize the clear distinction between the effectiveness testing currently underway, and the operational testing that can only be done later in the program development process. Until connectivity with the airlines is established, operational testing is not feasible. Consequently, many questions relating to operational effectiveness will necessarily remain unanswered until after the conclusion of testing.

www.dhs.gov
Appendix II: Comments from the Department of Homeland Security

Thank you again and we look forward to continued cooperation with the GAO on your reporting requirements for the House and Senate Appropriations Committees per P.L. 108.334 and any further reporting requested.

Sincerely,

[Signature]

Steven J. Pecinovsky
Acting Director, Departmental GAO/IG Liaison
Office of the Chief Financial Officer
GAO’s Mission

The Government Accountability Office, the audit, evaluation and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's Web site (www.gao.gov). Each weekday, GAO posts newly released reports, testimony, and correspondence on its Web site. To have GAO e-mail you a list of newly posted products every afternoon, go to www.gao.gov and select "Subscribe to Updates."

Order by Mail or Phone

The first copy of each printed report is free. Additional copies are $2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. Government Accountability Office
441 G Street NW, Room LM
Washington, D.C. 20548

To order by Phone: Voice: (202) 512-6000
TDD: (202) 512-2537
Fax: (202) 512-6061

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:
E-mail: fraudnet@gao.gov
Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Gloria Jarmon, Managing Director, JarmonG@gao.gov (202) 512-4400
U.S. Government Accountability Office, 441 G Street NW, Room 7125
Washington, D.C. 20548

Public Affairs

Paul Anderson, Managing Director, AndersonP1@gao.gov (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, D.C. 20548