Testimony

Before the Committee on Oversight and Government Reform, House of Representatives

AVIATION SECURITY

Improved Testing, Evaluation, and Performance Measurement Could Enhance Effectiveness

Statement of Jennifer Grover, Director, Homeland Security and Justice

GAO Highlights

AVIATION SECURITY

Improved Testing, Evaluation, and Performance Measurement Could Enhance Effectiveness

Why GAO Did This Study

Since the attacks of September 11, 2001 exposed vulnerabilities in the nation's aviation system, billions of dollars have been spent on a wide range of programs designed to enhance aviation security. Continuing fiscal pressure highlights the need for DHS's TSA, the primary federal agency responsible for aviation security, to allocate its finite resources for the greatest impact.

This testimony addresses the extent to which TSA has (1) evaluated the overall effectiveness of new technologies, programs, and processes using robust methods of testing and evaluation, (2) established performance measures that fully reflect program goals, and (3) used program data to identify opportunities for improvement. This statement is based on findings from GAO reports and testimonies issued from January 2013 through June 2015, with selected updates conducted from April 2015 through October 2015 to, among other things, determine progress made in implementing previous GAO recommendations. For prior work, GAO analyzed TSA policy documents and interviewed TSA officials, among other things. For the updates, GAO reviewed documents and followed up with TSA officials about actions to address GAO recommendations.

What GAO Recommends

GAO has previously made recommendations to DHS to strengthen TSA's aviation security programs. DHS generally agreed and has either addressed or has actions underway to address most of them. Consequently, GAO is not making new recommendations in this testimony.

What GAO Found

Evaluation of new technologies, programs, and processes. GAO has found that TSA has not consistently evaluated the overall effectiveness of new technologies before adopting them. For example, in March 2014, GAO found that TSA testing of certain Advanced Imaging Technology (AIT) systems—also referred to as full-body scanners—used to screen passengers at airports did not account for all factors affecting the systems. GAO reported that the effectiveness of AIT systems equipped with automated target recognition software (AIT-ATR)—which display anomalies on a generic passenger outline—relied on both the technology's capability to identify potential threat items and its operators' ability to resolve them. However, GAO found that TSA did not include operators' ability in determining overall AIT-ATR system performance. GAO recommended that TSA, in considering procurement of the next generation of AIT systems (AIT-2), measure system effectiveness based on the performance of both the technology and the screening personnel. The Department of Homeland Security (DHS) concurred and, in June 2015, TSA provided documentation showing that, while conducting operational testing of the AIT-2 system, the agency considered screening officer performance and measured AIT-2 system effectiveness based on both the performance of the AIT-2 technology and the screening officers who operate it. This should help TSA assess whether this screening system will meet mission needs and perform as intended.

Establishment of performance measures. GAO has found that TSA has not consistently established performance measures that fully reflect program goals. For example, in September 2014, GAO found that TSA's performance measures for Secure Flight—a passenger prescreening program—did not allow TSA to fully assess its progress toward achieving all of its goals. For example, one program goal was to accurately identify passengers on various watch lists, but TSA did not have measures to assess the extent of system matching errors, such as the extent to which Secure Flight is missing passengers who are actual matches to these lists. GAO recommended that TSA develop such measures. DHS concurred, and, as of April 2015, TSA was evaluating its current Secure Flight performance goals and measures and determining what new performance measures should be established to fully measure progress against program goals.

Use of program data. GAO has also reported on findings related to program data that TSA collected but had not analyzed, missing opportunities to refine and further improve TSA programs. For example, in January 2013, GAO reported that TSA collected and used key program data in support of its National Explosives Detection Canine Team Program, but could better analyze these data to identify program trends. For example, GAO found that in reviewing the results of certain covert tests, TSA did not analyze the results beyond the pass and fail rates, missing an opportunity to identify corrective actions. GAO recommended that TSA regularly analyze available data to identify program trends and areas that are working well and those in need of corrective action to guide program resources and activities. TSA concurred with GAO's recommendation and has taken actions to address this, including requiring analysis of the reasons for certain failed assessments.

Chairman Chaffetz, Ranking Member Cummings, and Members of the Committee:

I am pleased to be here today to discuss our past work examining the effectiveness of Transportation Security Administration (TSA) programs and technologies. It has been over 14 years since the attacks of September 11, 2001 exposed vulnerabilities in the nation's aviation system. Since then, billions of dollars have been spent on a wide range of programs designed to enhance aviation security. However, securing the nation's aviation operations remains a daunting task—with hundreds of airports, thousands of aircraft, and thousands of flights daily carrying millions of passengers and pieces of carry-on and checked baggage. According to TSA, the threat to civil aviation has not diminished—underscoring the need for effective aviation security programs. As the fiscal pressures facing the government continue, so too does the need for TSA to determine how to allocate its finite resources to have the greatest impact on addressing threats and strengthening the effectiveness of its programs and activities.

Over the past several years, TSA has taken numerous steps to strengthen aviation security. For example, TSA has deployed new screening technology intended to enhance passenger screening, developed processes and procedures to help ensure that individuals and their accessible property receive the appropriate level of screening, and established performance measures to assess progress toward achieving some program goals. However, we have identified opportunities to improve upon these efforts.

As requested, my testimony today identifies key issues that we have found to have adversely affected the effectiveness of TSA's aviation security investments and programs. Specifically, this testimony addresses the extent to which TSA has

(1) evaluated the overall effectiveness of new technologies, programs, and processes using robust methods of testing and evaluation;

(2) established performance measures that fully reflect program goals; and

(3) used program data to identify opportunities for improvement.

This statement is based on selected reports and testimonies issued by GAO from January 2013 through June 2015 related to TSA's efforts to

oversee its aviation security measures.[1] In addition, this statement is based on selected updates conducted from April 2015 through October 2015 related to the current status of the Secure Flight and Behavior Detection and Analysis programs and Managed Inclusion process, and progress made in implementing previous GAO recommendations.[2] For our past work, we reviewed applicable laws, regulations, and agency and departmental policies and TSA program documents; decision memorandums; results from screener performance reviews and testing of Advanced Imaging Technology (AIT), also referred to as full-body scanners; and other documents. We also visited airports—four for our Behavior Detection work, six for our Managed Inclusion work, and nine for our Secure Flight work—which we selected based on a variety of factors, such as volume of passengers screened and geographic dispersion, and interviewed Department of Homeland Security (DHS), TSA, and Federal Bureau of Investigation officials, among other things. Further details on the scope and methodology for the previously issued reports and testimonies are available within each published product. For the updates, we reviewed documents and followed up with TSA officials related to the actions taken to address our recommendations. We conducted the work on which this statement is based in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

---

[1]See GAO, *TSA Explosives Detection Canine Program: Actions Needed to Analyze Data and Ensure Canine Teams Are Effectively Utilized,* GAO-13-239 (Washington, D.C.: Jan. 31, 2013); *Aviation Security: TSA Should Limit Future Funding for Behavior Detection Activities,* GAO-14-159 (Washington, D.C.: Nov. 8, 2013); *Advanced Imaging Technology: TSA Needs Additional Information before Procuring Next-Generation Systems,* GAO-14-357 (Washington, D.C.: Mar. 31, 2014); *Secure Flight: TSA Should Take Additional Steps to Determine Program Effectiveness,* GAO-14-531 (Washington, D.C.: Sept. 9, 2014); *Aviation Security: Rapid Growth in Expedited Passenger Screening Highlights Need to Plan Effective Security Assessments,* GAO-15-150 (Washington, D.C.: Dec. 12, 2014); and *Aviation Security: TSA Has Taken Steps to Improve Oversight of Key Programs, but Additional Actions Are Needed,* GAO-15-678T, (Washington, D.C.: June 9, 2015).

[2]Secure Flight is a passenger prescreening program that matches passenger information against federal government watch lists and other information to assign each passenger to a risk category. Managed Inclusion is a process that TSA uses to determine passengers' eligibility for expedited screening at some passenger screening checkpoints. TSA's Behavior Detection and Analysis program uses trained officers to identify passenger behaviors indicative of stress, fear, or deception and refer passengers meeting certain criteria for additional screening of their persons and carry-on baggage.

Background

The Aviation and Transportation Security Act (ATSA) established TSA as the primary federal agency with responsibility for securing the nation's civil aviation system.[3] This responsibility includes the screening of all passengers and property transported from and within the United States by commercial passenger aircraft.[4] In accordance with ATSA, all passengers, their accessible property, and their checked baggage are to be screened pursuant to TSA-established procedures at the more than 450 airports at which TSA performs, or oversees the performance of, security screening operations. These procedures generally provide, among other things, that passengers pass through security checkpoints where their person, identification documents, and accessible property are checked by screening personnel. The following are some of TSA's transportation security technologies, processes, and programs.

AIT systems: According to TSA officials, AIT systems provide enhanced security benefits compared with those of walk-through metal detectors by identifying nonmetallic objects and liquids. Following the deployment of AIT, the public and others raised privacy concerns because AIT systems produced images of passengers' bodies that image operators analyzed to identify objects or anomalies that could pose a threat to an aircraft or to the traveling public. To mitigate those concerns, TSA began installing automated target recognition (ATR) software on deployed AIT systems in July 2011.[5] AIT systems equipped with ATR (AIT-ATR) automatically interpret the image and display anomalies on a generic outline of a passenger instead of displaying images of actual passenger bodies. Screening officers use the generic image of a passenger to identify and resolve anomalies on-site in the presence of the passenger.

Expedited Screening and TSA's Managed Inclusion Process: TSA Pre✓™—TSA's expedited screening program—is intended to allow TSA to devote more time and resources at the airport to screening the passengers TSA determined to be higher or unknown risk while providing expedited screening to those passengers determined to pose a lower risk

---

[3]Pub. L. No. 107-71, 115 Stat. 597 (2001).

[4]See 49 U.S.C. § 44901. For purposes of this testimony, "commercial passenger aircraft" refers to U.S.- or foreign-flagged air carriers operating under TSA-approved security programs with regularly scheduled passenger operations to or from a U.S. airport.

[5]See Pub. L. No. 112-95, § 826, 126 Stat. 11, 132-33 (2012) (codified at 49 U.S.C. § 44901(l)) (requiring, in general, that TSA ensure that all AIT systems used to screen passengers are equipped with ATR software).

to the aviation system. To assess whether a passenger is eligible for expedited screening, TSA considers, in general, (1) inclusion on an approved TSA Pre✓™ list of known travelers;[6] (2) results from the automated TSA Pre✓™ risk assessments of all passengers;[7] and (3) real-time threat assessments of passengers, known as Managed Inclusion, conducted at airport checkpoints. Through its Managed Inclusion process, TSA has utilized a combination of security measures, including behavior detection officers (BDO) and passenger screening canine teams at the checkpoint to identify passengers suitable for expedited screening.[8]

Behavior Detection and Analysis program: TSA's Behavior Detection and Analysis program, formerly known as the Screening of Passengers by Observation Techniques (SPOT) program, is intended to identify persons who may pose a risk to aviation security. Through these behavior detection activities, TSA's BDOs are to identify passenger behaviors indicative of stress, fear, or deception and refer passengers meeting certain criteria for additional screening of their persons and carry-on baggage. During this referral screening, if passengers exhibit additional such behaviors, or if other events occur, such as the discovery of a suspected fraudulent document, BDOs are to refer these passengers to a law enforcement officer for further investigation, which could result in an arrest, among other outcomes.

Secure Flight: Since TSA began implementing Secure Flight in 2009, the passenger prescreening program has changed from a program that identifies passengers as high risk solely by matching them against federal government watch lists—for example, the No Fly List, comprised of individuals who should be precluded from boarding an aircraft, and the Selectee List, comprised of individuals who should receive enhanced

---

[6]This includes individuals who have been preapproved as low-risk travelers through application to the TSA Pre✓™ program as well as individuals whom TSA has determined to be low risk by virtue of their membership in a specific group, such as active duty military members.

[7]TSA uses passenger travel-related data to assess passenger risk and assign passengers scores to identify low-risk passengers eligible for expedited screening for a specific flight prior to the passengers' arrival at the airport.

[8]BDOs identify passenger behaviors indicative of stress, fear, or deception and refer passengers meeting certain criteria for additional screening of their persons and carry-on baggage. TSA uses passenger screening canine teams comprised of a canine paired with a handler to deter and detect the use of explosive devices.

screening at the passenger security checkpoint—to one that uses additional lists and risk-based criteria to assign passengers to a risk category: high risk, low risk, or unknown risk.[9] In 2010, following the December 2009 attempted attack on a U.S.-bound flight, which exposed gaps in how agencies used watch lists to screen individuals, TSA began using risk-based criteria to create additional lists for Secure Flight screening. These lists are composed of high-risk passengers who may not be in the Terrorist Screening Database (TSDB)—the U.S. government's consolidated watch list of known or suspected terrorists—but whom TSA has determined should be subject to enhanced screening procedures.[10] Further, in 2011, TSA began screening passengers against additional identities in the TSDB that are not included on the No Fly or Selectee Lists. In addition, as part of TSA Pre✓™, TSA began screening against several new lists of preapproved low-risk travelers. TSA also began conducting TSA Pre✓™ risk assessments—an activity distinct from matching against watch lists—that use the Secure Flight system to assign passengers scores based upon their travel-related data, for the purpose of identifying them as low risk for a specific flight.

National Explosives Detection Canine Team Program: One of TSA's security layers is its National Explosives Detection Canine Team Program (NEDCTP), composed of over 800 explosives detection canine teams—a canine paired with a handler—aimed at deterring and detecting the use of explosive devices in the U.S. transportation system. Through NEDCTP, TSA trains, deploys, and certifies explosives detection canine teams. TSA deploys the teams to screen passengers and air cargo at airports and other transportation modes, including mass transit.

---

[9]The No Fly and Selectee Lists are subsets of the Terrorist Screening Database—the U.S. government's consolidated watch list of known or suspected terrorists.

[10]Standard screening typically includes passing through a walk-through metal detector or AIT system, which identifies objects or anomalies on the outside of the body, and X-ray screening for the passenger's accessible property. In general, enhanced screening includes, in addition to the procedures applied during a typical standard screening experience, a pat-down and an explosives trace detection or physical search of the interior of the passenger's accessible property, electronics, and footwear.

TSA Has Not Consistently Evaluated the Overall Effectiveness of New Technologies, Programs, and Processes Using Robust Methods of Testing and Evaluation

In our 2014 reviews of TSA's AIT-ATR systems and Managed Inclusion process, we found that TSA had conducted some testing before adopting the new technology and process, but it had not fully demonstrated their effectiveness. We have also previously reported on challenges TSA has faced in designing studies and protocols to test the effectiveness of security systems and programs in accordance with established methodological practices, such as in the case of our 2013 review of TSA's behavior detection activities. TSA has since taken steps to more comprehensively test the effectiveness of the next generation of AIT, known as AIT-2, and further test aspects of the Managed Inclusion process.

With regard to the AIT-ATR system, in March 2014, we reported that, according to TSA officials, checkpoint security is a function of technology, people, and the processes that govern them; however, we found that TSA did not include each of those factors in determining overall AIT-ATR system performance.[11] Specifically, we found that TSA evaluated the technology's performance in the laboratory to determine system effectiveness. Laboratory test results provide important insights but do not accurately reflect how well the technology will perform in the field with actual human operators. Additionally, we found that TSA did not assess how alarms are resolved by considering how the technology, people, and processes function collectively as an entire system when determining AIT-ATR system performance. AIT-ATR system effectiveness relies on both the technology's capability to identify threat items and its operators to resolve those threat items.

Given that TSA was seeking to procure the second generation of AIT systems, known as AIT-2, we reported that DHS and TSA would be hampered in their ability to ensure that future AIT systems meet mission needs and perform as intended at airports unless TSA evaluated system effectiveness based on both the performance of the AIT-2 technology and screening officers who operate the technology. According to best practices related to federal acquisitions, technologies should be demonstrated to work in their intended environment. We recommended that TSA measure system effectiveness based on the performance of the AIT-2 technology and screening officers who operate the technology while taking into account current processes and deployment strategies. TSA concurred and has addressed this recommendation. Specifically, in

---

[11]GAO-14-357.

June 2015, TSA provided documentation showing that, while conducting operational testing of the AIT-2 system, the agency considered screening officer performance and measured AIT-2 system effectiveness based on both the performance of the AIT-2 technology and the screening officers who operate it. This should help TSA assess whether this screening system will meet mission needs and perform as intended.

With regard to the Managed Inclusion process, in December 2014, we reported that TSA had tested the security effectiveness of the individual components of the Managed Inclusion process, but had not tested the overall effectiveness of the Managed Inclusion process as it functions as a whole.[12] According to TSA officials, TSA tested the security effectiveness of the individual components being used in the Managed Inclusion process at the time—such as BDOs, passenger screening canine teams, and explosives trace detection (ETD) devices—before implementing Managed Inclusion, and TSA determined that each layer alone provides an effective level of security.[13] However, in our prior body of work, we identified challenges in several of the layers used in the Managed Inclusion process, raising questions regarding their effectiveness.[14] Further, as of the time of our report, TSA officials stated that they had not yet tested the security effectiveness of the Managed Inclusion process as it functions as a whole. TSA officials explained that they had been planning to test the process as a whole and estimated that such testing would begin in October 2014 and would take 12 to 18 months to complete. However, TSA could not provide us with specifics or a plan or documentation showing how the testing was to be conducted, the locations where it was to occur, how those locations were to be selected, or the timeframes for conducting testing at each location. In

---

[12]GAO-15-150.

[13]In September 2015, TSA officials stated that they had suspended the use of one type of Managed Inclusion that used a combination of BDOs and ETD technology but will continue to operate Managed Inclusion with BDOs and passenger screening canines.

[14]For example, in our November 2013 report on TSA's Behavior Detection and Analysis program, we found that although TSA had taken several positive steps to validate the scientific basis and strengthen program management of its Behavior Detection and Analysis program, TSA had not demonstrated that behavioral indicators can be used to reliably and effectively identify passengers who may pose a threat to aviation security. See GAO-14-159; *Explosives Detection Canines: TSA Has Taken Steps to Analyze Canine Team Data and Assess the Effectiveness of Passenger Screening Canines,* GAO-14-695T (Washington, D.C.: June 24, 2014); and *Aviation Security: TSA Has Enhanced Its Explosives Detection Requirements for Checked Baggage, but Additional Screening Actions Are Needed,* GAO-11-740 (Washington, D.C.: July 11, 2011).

general, evaluations are most likely to be successful when key steps are addressed during design, including defining research questions appropriate to the scope of the evaluation, and selecting appropriate measures and study approaches that will lead to valid conclusions. As a result, we recommended that to ensure TSA's planned testing yields reliable results, the TSA Administrator take steps to ensure that TSA's testing of the Managed Inclusion process adheres to established evaluation design practices.[15]

DHS concurred with our recommendation and has taken some initial steps toward addressing it. Specifically, in August 2015, TSA officials provided us with a testing schedule for additional testing of canine teams and BDOs—two of the security layers used in the Managed Inclusion process—and stated that they had plans to ensure that the tests adhere to recognized test and evaluation protocols. However, TSA has not provided documents explaining how it plans to evaluate the Managed Inclusion process as a whole and how this evaluation will adhere to established evaluation practices. These documents would need to constitute a research plan specifically tailored to evaluating Managed Inclusion and include specifics such as the types of data to be collected, the methodology for collecting and analyzing the data, and the steps TSA plans to take to help ensure the study will isolate the security effects of the Managed Inclusion process itself and rule out plausible alternative explanations for study results.

Further, in November 2013, we found that a 2011 DHS study conducted to validate SPOT's behavioral indicators did not demonstrate their effectiveness because of study limitations, including the use of unreliable data.[16] We concluded that the usefulness of DHS's April 2011 validation study was limited, in part because the data the study used to examine the extent to which the SPOT behavioral indicators led to correct screening decisions at security checkpoints were from the SPOT database that we had previously found in May 2010 to have several weaknesses, and thus were potentially unreliable.[17] Specifically, in May 2010, we assessed the

---

[15]GAO, *Designing Evaluations: 2012 Revision,* GAO-12-208G (Washington, D.C.: Jan. 31 2012).

[16]GAO-14-159.

[17]GAO, *Aviation Security: Efforts to Validate TSA's Screening Behavior Detection Program Underway, but Opportunities Exist to Strengthen Validation and Address Operational Challenges,* GAO-10-763 (Washington, D.C.: May 20, 2010).

reliability of the SPOT database and concluded that the database lacked controls to help ensure the completeness and accuracy of the data, such as computerized edit checks to review the format, existence, and reasonableness of data. In that report, we also found, among other things, that BDOs could not record all behaviors observed in the SPOT database because the database limited entry to eight behaviors, six signs of deception, and four types of serious prohibited items per passenger referred for additional screening. At that time, BDOs were trained to identify 94 signs of stress, fear, and deception, or other related indicators.[18] In May 2010, we recommended that TSA make changes to ensure the quality of SPOT referral data, and TSA subsequently made changes to the SPOT database. However, we found in our 2013 report that DHS's validation study used data that were collected from 2006 through 2010, prior to TSA's improvements to the SPOT database. As a result, we determined that the data used in the SPOT validation study were not reliable enough for TSA to use in conducting a statistical analysis of the association between the indicators and high-risk passenger outcomes.[19] Because the study used unreliable data, its conclusions regarding the use of the SPOT behavioral indicators for passenger screening were questionable and did not support the conclusion that the indicators can or cannot be used to identify threats to aviation security.

Due to these and other methodological issues we found in DHS's validation study, we recommended in November 2013 that the Secretary of Homeland Security direct TSA to limit future funding support for the agency's behavior detection activities until TSA can provide scientifically validated evidence that demonstrates that behavioral indicators can be

---

[18]The 2011 SPOT standard operating procedures listed 94 signs of stress, fear, and deception, or other related indicators that BDOs are to look for, each of which is assigned a certain number of points.

[19]The SPOT validation study analyzed data collected from 2006 to 2010 to determine the extent to which the indicators could identify high-risk passengers defined as passengers who (1) possessed fraudulent documents, (2) possessed serious prohibited or illegal items, (3) were arrested by a law enforcement officer, or (4) any combination of the first three measures. The validation study reported that 14 of the 41 SPOT behavioral indicators were positively and significantly related to one or more of the study outcomes. These outcome measures were developed for the validation study. Possession of fraudulent documents is a subset of possession of serious prohibited or illegal items. According to the validation study, the possession of fraudulent documents was studied independently as an outcome measure, since it was the largest class of serious prohibited or illegal items.

used to identify passengers who may pose a threat to aviation security. DHS did not concur with our recommendation, in part because it disagreed with our analysis of TSA's behavioral indicators. In January 2015, TSA provided documentation describing its plans to enhance its behavioral-based screening program, including the development of revised behavioral indicators and new protocols for their use. In October 2015, TSA officials told us they were in the process of pilot testing the new protocols in the airport environment and expect to complete the tests by February 2016. Officials stated that TSA plans to make a determination about whether the new protocols are ready for further testing, including an operational test to determine the protocol's effectiveness, at that time. Further, TSA officials estimated that the operational test may begin in the summer of 2016, but they did not have an estimated completion date because the behavior detection covert test methodology had not yet been developed and the threat inject methods had not yet been deemed sufficiently mature to test effectiveness.[20] Until TSA completes its planned tests and study on the use of the new protocols and provides the scientifically validated evidence of effectiveness, such as successful operational testing, the agency continues to fund activities that have not been determined to be effective.

TSA Has Not Consistently Established Performance Measures That Fully Reflect Program Goals

In 2014, we reported on two instances in which TSA's performance measures made it difficult to assess TSA's performance in meeting its goals. First, we found that TSA did not have adequate performance measures for all Secure Flight program goals and second, we found that TSA tracked performance information on the expedited screening program that did not link to program goals.

In September 2014, we found that Secure Flight had established program goals that reflect new program functions since implementation began in 2009 to identify additional types of high-risk and also low-risk passengers; however, the program performance measures in place at that time did not allow TSA to fully assess its progress toward achieving all of its goals.[21] For example, one program goal was to accurately identify passengers on

---

[20]In October 2015, TSA officials stated that they also plan to conduct a study on the use of the new protocols at 50 airports to examine disparity questions regarding racial, ethnicity, and religious garb demographics. According to these officials, this study, which should begin at airports as the new protocols are implemented and require 12 to 15 months of data collection, is not expected to be completed until 2018.

[21]GAO-14-531.

various watch lists. To assess performance toward this goal, Secure Flight collected various types of data, including the number of passengers TSA identifies as matches to high- and low-risk lists, but did not have measures to assess the extent of system matching errors—for example, the extent to which Secure Flight is missing passengers who are actual matches to these lists. We concluded that additional measures that address key performance aspects related to program goals, and that clearly identify the activities necessary to achieve goals, in accordance with the Government Performance and Results Act, would allow TSA to more fully assess progress toward its goals.[22] Therefore, we recommended that TSA develop such measures, and ensure these measures clearly identify the activities necessary to achieve progress toward the goal. DHS concurred with our recommendation, and, according to TSA officials, as of April 2015, TSA's Office of Intelligence and Analysis was evaluating its current Secure Flight performance goals and measures and determining what new performance measures should be established to fully measure progress against program goals. Establishing additional performance measures that adequately indicate progress toward goals would allow Secure Flight to more fully assess the extent to which it is meeting program goals.

Further, in December 2014, we reported that TSA's performance measure for assessing its expedited screening program did not accurately link to the program's goals—to ensure: (1) that 25 percent of air passengers were eligible for expedited screening by the end of calendar year 2013, and (2) that 50 percent of passengers were eligible for expedited screening by the end of calendar year 2014.[23] According to TSA documents, TSA uses one measure—the total number of air passengers screened daily using expedited screening as a percentage of the total number of passengers screened daily—to assess progress

---

[22]Government Performance and Results Act (GPRA) of 1993, Pub. L. No. 103-62, 107 Stat. 285, as updated by the GPRA Modernization Act of 2010, Pub. L. No. 111-352, 124 Stat. 3866 (2011).

[23]GAO-15-150.

toward these goals.[24] TSA collects data for this measure by reporting, not the number of passengers designated as eligible for expedited screening, but the number of passengers who actually receive such screening. However, because expedited screening is voluntary, not all passengers who are eligible necessarily use expedited screening. For example, a passenger may be traveling with a group in which not all passengers in the group are eligible for expedited screening, so the passenger may choose to forgo expedited screening. As a result, the information that TSA is reporting to show that it is meeting its goal may be understated and inaccurate.

TSA's Chief Risk Officer agreed that the goals and the measure are not linked, but said that tracking actual screening data rather than eligibility data presents a more accurate picture of the expedited screening program performance. While we agreed that tracking actual screening data may provide insights about expedited screening program performance, we reported that ensuring goals and measures are aligned is important to provide more accurate performance measurement data to guide program performance and to identify potential areas for improvement. Best practices regarding the key attributes of successful performance measurement state that performance measures should link and align with agency-wide goals and the mission should be clearly communicated throughout the organization. We recommended that the TSA Administrator align TSA's expedited screening performance goals and measures to ensure that TSA, as well as lawmakers, has accurate information by which to measure the performance of its expedited screening programs. TSA has not yet addressed this recommendation.

---

[24]In addition, TSA used seven other measures to track other screening activities daily and reported the performance results of these measures to airport federal security directors (FSD)—TSA officials with overall responsibility for security operations at one or more airports—to keep them apprised of this progress. According to TSA officials, these additional measures are intended to be internal program measures and are not published outside of TSA. The daily report includes targets that FSDs are to meet daily for each of these measures, and if these targets are not met for a period of 4 weeks, the FSD is to develop an action plan to show how the airport plans to work to meet the targets.

TSA Has Not Consistently Used Program Data to Identify Opportunities for Improvement

We have also reported on findings related to program data—such as canine program assessment data and Secure Flight screening error data—that TSA collected but had not analyzed, missing opportunities to refine and further improve TSA programs.

In January 2013, we reported that TSA collected and used key canine program data in support of its NEDCTP program, but could better analyze these data to identify program trends.[25] For example, we found that in reviewing short notice assessments (covert tests), TSA did not analyze the results beyond the pass and fail rates. Therefore, TSA was missing an opportunity to determine, for example, if there were any search areas or types of explosives in which canine teams were more effective compared with others, and what, if any, training may be needed to mitigate deficiencies. *Standards for Internal Control in the Federal Government* calls for agencies to ensure that ongoing monitoring occurs during the course of normal operations to help evaluate program effectiveness.[26] We recommended that TSA regularly analyze available data to identify program trends and areas that are working well and those in need of corrective action to guide program resources and activities. TSA concurred with our recommendation and has taken actions that address our recommendation. For example, in the event a canine team fails a short notice assessment, TSA now requires that canine team supervisors complete an analysis of the team's training records to identify an explanation for the failure.

Further, in September 2014, we reported that TSA has processes in place to implement Secure Flight screening determinations at airport checkpoints, but could evaluate available data on screening errors to identify corrective actions.[27] TSA information from May 2012 through February 2014 that we assessed indicated that screening personnel had made errors in implementing Secure Flight determinations at the checkpoint. TSA officials we spoke with at five of the nine airports where we conducted interviews stated that they conduct after-action reviews of

---

[25]GAO-13-239.

[26]GAO, *Internal Control: Standards for Internal Control in the Federal Government,* GAO/AIMD-00-21.3.1 (Washington, D.C.: November 1999).

[27]See GAO-14-531. Screening personnel at airport checkpoints are primarily responsible for ensuring that passengers receive a level of screening that corresponds to the level of risk determined by Secure Flight by verifying passengers' identities and identifying passengers' screening designations.

screening errors at the checkpoint and have used these reviews to take action to address the root causes of those errors. However, we found that TSA did not have a systematic process for evaluating the root causes of these screening errors at the checkpoint across airports, which could allow TSA to identify trends across airports and target nationwide efforts to address these issues. Consistent with *Standards for Internal Control in the Federal Government,* we recommended in September 2014 that TSA develop a process for evaluating the root causes of screening errors at the checkpoint and then implement corrective measures to address those causes.[28] DHS concurred with our recommendation and has taken actions to address them. Specifically, TSA provided us with documentation of its analysis of screening errors that occurred over a 3 month period—June 2015 through early September 2015—and the root causes of those errors. Additionally, in September 2015, TSA made changes to its screening procedures to address the root causes of errors identified through its analysis.

Chairman Chaffetz, Ranking Member Cummings, and Members of the Committee, this completes my prepared statement. I would be pleased to respond to any questions that you may have at this time.

GAO Contacts and Staff Acknowledgments

For questions about this statement, please contact Jennifer Grover at (202) 512-7141 or groverj@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this statement. Individuals making key contributions to this statement include Maria Strudwick (Assistant Director), Claudia Becker, Michele Fejfar, Susan Hsu, and Tom Lombardi. Key contributors for the previous work that this testimony is based on are listed in each product.

---

[28] GAO/AIMD-00-21.3.1.