United States Government Accountability Office

Report to Congressional Committees

March 2021

F-35 JOINT STRIKE FIGHTER

DOD Needs to Update Modernization Schedule and Improve Data on Software Development

A Century of Non-Partisan Fact-Based Work

GAO-21-226

GAO@100

Highlights

Highlights of GAO-21-226, a report to congressional committees

March 2021

F-35 JOINT STRIKE FIGHTER

DOD Needs to Update Modernization Schedule and Improve Data on Software Development

Why GAO Did This Study

The F-35 Lightning II Joint Strike Fighter program remains DOD's most expensive weapon system program. DOD is 3 years into a development effort that is loosely based on Agile software development processes to modernize the F-35 aircraft's capabilities. With this approach, DOD intends to incrementally develop, test, and deliver small groups of new capabilities every 6 months. Congress included provisions in two statutes for GAO to review the F-35 program.

This report addresses the F-35 operational testing status, DOD's Block 4 modernization development schedule, and how the F-35 program office implements key practices for evaluating Agile software development progress. To assess cost and schedule concerns identified in prior years, GAO selected three key practices that focus on evaluating Agile software development progress. GAO reviewed DOD and contractor documentation and interviewed DOD officials and contractor representatives.

What GAO Recommends

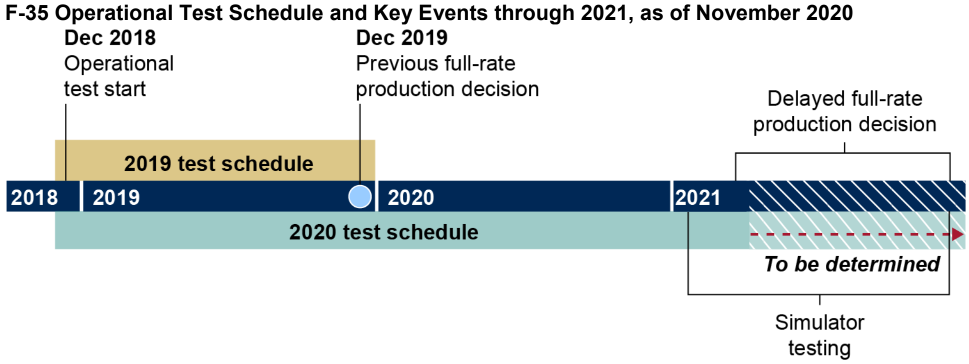
GAO is making three recommendations to DOD, including that it update its modernization schedule to reflect achievable time frames, identify and implement tools to enable automated data collection on software development performance, and set software quality performance targets. DOD agreed with GAO's recommendations.

View GAO-21-226. For more information, contact Jon Ludwigson at (202) 512-4841 or ludwigsonj@gao.gov.

What GAO Found

The Department of Defense (DOD) delayed the completion of key testing until problems with the F-35 aircraft simulator are resolved, which GAO also reported last year, and will again delay its full-rate production decision. In August 2020, the program office determined the aircraft simulator—to be used to replicate complex test scenarios that could not be accomplished in real-world environment testing—did not fully represent F-35 capabilities and could not be used for further testing until fixed. Since then, program officials have been developing a new plan to ensure the simulator works as intended. Until they finalize the plan and fix the simulator, the next production milestone date—which would formally authorize DOD's transition from development to full production—remains undetermined (see figure).

F-35 Operational Test Schedule and Key Events through 2021, as of November 2020



Source: GAO analysis of Department of Defense data. | GAO-21-226

DOD is now in its third year of its modernization effort, known as Block 4, to upgrade the hardware and software of the aircraft. While DOD added another year to the schedule, GAO found the remaining development time frame is not achievable. The program routinely underestimated the amount of work needed to develop Block 4 capabilities, which has resulted in delays, and has not reflected historical performance into its remaining work schedule. Unless the F-35 program accounts for historical performance in the schedule estimates, the Block 4 schedule will continue to exceed estimated time frames and stakeholders will lack reliable information on when capabilities will be delivered.

GAO found the F-35 program office collects data on many Block 4 software development metrics, a key practice from GAO's Agile Assessment Guide, but has not met two other key practices for monitoring software development progress. Specifically, the F-35 program office has not implemented tools to enable automated data collection on software development performance, a key practice. The program's primary reliance on the contractor's monthly reports, often based on older data, has hindered program officials' timely decision-making. The program office has also not set software quality performance targets, inconsistent with another key practice. Without these targets, the program office is less able to assess whether the contractor has met acceptable quality performance levels.

Contents

Figures

Abbreviations

ALIS	Autonomic Logistics Information System
C2D2	Continuous Capability Development and Delivery
COVID-19	Coronavirus Disease 2019
DOD	Department of Defense
DOT&E	Director of Operational Test and Evaluation
NDAA	National Defense Authorization Act
ODIN	Operational Data Integrated Network
TR-2	Technology Refresh 2
TR-3	Technology Refresh 3

441 G St. N.W.
Washington, DC 20548

March 18, 2021

Congressional Committees

The F-35 Lightning II Joint Strike Fighter program is a family of fifth-generation strike fighter aircraft that integrates low-observable (stealth) technology with advanced sensors and computer networking capabilities for the United States Air Force, Marine Corps, and Navy, as well as seven international partners. The program aims to procure 2,470 F-35s to replace several other aircraft used by the Air Force, Navy, and Marine Corps to perform a wide range of missions. To date, the program has delivered over 600 aircraft to the U.S. services, allied partners, and foreign military sales customers.

The Department of Defense (DOD) is now in the third year of a $14 billion modernization effort—known as Block 4—to upgrade the hardware and software systems of the F-35. DOD intends for Block 4 to modernize the aircraft and address new threats that have emerged since the aircraft's original requirements were established in 2000. DOD is using a different development approach for Block 4, referred to as Continuous Capability Development and Delivery (C2D2), which is loosely based on Agile software development processes. With this approach, DOD intends to deliver capabilities to the warfighter faster than it did during the original development program.

The program wrapped up development of the F-35's original capabilities in 2018 and is undergoing operational testing to verify that the aircraft adequately provide those baseline capabilities. According to program officials, prior to October 2020, the program expected to complete this testing in January 2021 and hold a full-rate production decision—which would formally authorize DOD's transition from development to full production—in March 2021. As the program moves toward completing this testing and evaluating the results, it still faces risks ahead of the full-rate production decision. We reported on these and other program risks in the past and made recommendations for improvement. DOD has taken action to address some, but not all, of our recommendations. For a comprehensive list of our recommendations and a summary of DOD's actions in response, see appendix I. In addition, a list of related GAO products is included at the end of the report.

This report fulfills two mandates. First, the National Defense Authorization Act (NDAA) for Fiscal Year 2015 included a provision for GAO to submit a report on the F-35 program's production and Block 4 progress within 30 days of the President's budget submission for Fiscal Years 2021-2025. This is our sixth report under that provision.[1] Second, the NDAA for Fiscal Year 2020 includes a provision for GAO to submit a report on the F-35 program's production and Block 4 progress within 30 days of the President's budget submission for Fiscal Years 2021-2025. This is our second report under that provision.[2]

In this report, we (1) identify and describe any remaining risks with completing operational testing for the baseline program ahead of the next production milestone decision, and the steps DOD took to mitigate them; (2) assess DOD's progress in developing and delivering Block 4 modernization capabilities and the program's efforts to address any remaining risks; and (3) determine the extent to which the F-35 program office is addressing key selected Agile software development practices as it implements Block 4 development.

- To identify and describe the remaining risks with the baseline program's operational testing completion ahead of the next production milestone decision, and the steps DOD took to mitigate them, we reviewed test events, schedules, program briefings, and DOD briefings. We also discussed key aspects of F-35 operational test progress with program management, contractor representatives, test pilots, and DOD test officials.
- To assess DOD's progress in developing and delivering Block 4 modernization capabilities and remaining risks, we reviewed program documentation, including cost and schedule estimates for capability development and testing, and interviewed DOD officials and contractor representatives.

[1] GAO- *F-35 Joint Strike Fighter: Actions Needed to Address Manufacturing and Modernization Risks,* GAO-20-339 (Washington, D.C.: May 12, 2020); *F-35 Joint Strike Fighter: Action Needed to Improve Reliability and Prepare for Modernization Efforts*, GAO-19-341 (Washington, D.C.: Apr. 29, 2019); *F-35 Joint Strike Fighter: Development Is Nearly Complete, but Deficiencies Found in Testing Need to Be Resolved,* GAO-18-321 (Washington, D.C.: June 5, 2018); *F-35 Joint Strike Fighter: DOD Needs to Complete Developmental Testing Before Making Significant New Investments*, GAO-17-351 (Washington, D.C.: Apr. 24, 2017); and *F-35 Joint Strike Fighter: Continued Oversight Needed as Program Plans to Begin Development of New Capabilities*, GAO-16-390 (Washington, D.C.: Apr. 14, 2016).

[2] GAO-20-339.

- To determine the extent that the F-35 program office addresses selected key practices for Agile software development, we first analyzed GAO's Agile Assessment Guide, the Defense Innovation Board "Software is Never Done" report, Defense Acquisition University, DOD's "Contracting Considerations for Agile Solutions: Key Agile Concepts and Sample Work Statement Language," and DOD's Operation of the Software Acquisition Pathway.[3] From those documents, we identified three key practices that would enable us to assess how the program uses Agile software development data to manage cost and schedule concerns that we identified in prior reports. These three key practices focus on evaluating Agile software development progress: (1) tracking metrics, (2) automating real-time data collection, and (3) establishing performance targets. We then analyzed F-35 Block 4 software development metrics and related documents and compared the F-35 program office's practices to these three key practices. We also interviewed program office officials knowledgeable on metrics for assessing software development, DOD officials, and contractor representatives.

We determined that all the data we used were sufficiently reliable for the purposes of responding to our reporting objectives. Appendix II contains a more detailed description of our scope and methodology.

We conducted this performance audit from March 2020 to March 2021 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.

Background

DOD started the F-35 program in 2001 to develop a fifth-generation fighter aircraft intended to replace a range of aging aircraft in the U.S. military services' inventories and to provide enhanced capabilities to warfighters that capitalized on technological innovations. Among other capabilities, the program designed the F-35 aircraft to be difficult to

[3]GAO, *GAO Agile Assessment Guide: Best Practices for Agile Adoption and Implementation,* GAO-20-590G (Washington, D.C.: Sept. 28, 2020); Defense Innovation Board, *Software is Never Done: Refactoring the Acquisition Code For Competitive Advantage* (May 2019); Acquisition and Sustainment, DOD, *Contracting Considerations for Agile Solutions: Key Agile Concepts and Sample Work Statement Language,* Version 1.0 (November 2019); and DOD Instruction 5000.87 Operation of the Software Acquisition Pathway (Oct. 2, 2020).

observe using radar and included sensors that can provide insights into potential targets and provide other warfighting information. Lockheed Martin is the prime contractor for the F-35 airframe and is responsible for integrating the engine into the airframe. Pratt & Whitney is contractor for the engine, also known as the F135.[4]

The program is producing and delivering three variants of the F-35 aircraft:

- the F-35A conventional takeoff and landing variant for the Air Force,
- the F-35B short takeoff and vertical landing variant for the Marine Corps, and
- the F-35C carrier-suitable variant for the Marine Corps and the Navy.

The characteristics of the services' variants are similar, but each variant also has unique operating requirements. For example, the Marine Corps requires that the F-35B be capable of operating from aircraft carriers, amphibious ships, as well as both main and austere operating bases. Figure 1 shows an F-35B preparing for flight.

[4]The engines are purchased by the government directly from Pratt & Whitney and delivered as government furnished equipment to Lockheed Martin for integration into the airframes during production.

Figure 1: An F-35B Exercising Its Short Takeoff and Vertical Landing Capability on the USS America

Source: © Lockheed Martin. | GAO-21-226

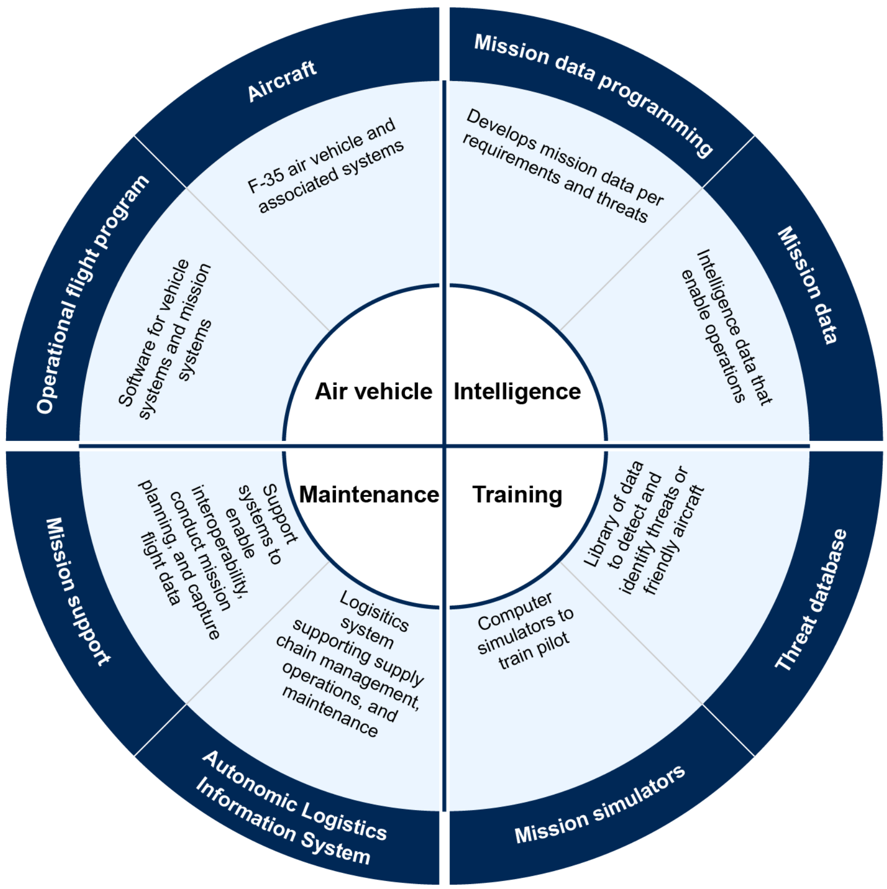
DOD leads the F-35 program but the program also involves several allied partner countries.[5] Companies in these countries also support aircraft production. In July 2019, DOD decided to remove Turkey from the development program due to its government's decision to procure Russian-made radar systems. Consequently, the F-35 program office and the prime contractors have identified and are contracting with alternative suppliers to produce the 1,005 parts that are currently made in Turkey.

While DOD plans to purchase 2,470 aircraft for the U.S. services, the F-35 program is acquiring more than just aircraft. The complete F-35 air system has eight elements, including training and maintenance systems. For the F-35 aircraft to be fully operational, capabilities associated with each element of the air system have to be developed and fielded in sync

[5]Seven partner nations—United Kingdom, Italy, Netherlands, Canada, Australia, Denmark, and Norway—contribute to the F-35 development, production, and sustainment. In addition, the program currently has six foreign military sales customers: Israel, Korea, Japan, Belgium, Poland, and Singapore. According to program officials, nine other countries are at various stages of foreign military sales consideration.

with the aircraft. Figure 2 shows the eight elements that make up the entire F-35 air system and how they support the aircraft.

Figure 2: The Eight Elements of the F-35 Air System



Source: GAO interpretation of Department of Defense data. | GAO-21-226

Note: The program is transitioning from the Automated Logistics Information System to the Operational Data Integrated Network.

One of these elements, the Automated Logistics Information System (ALIS), has not lived up to the program's expectations. DOD intends ALIS to provide the necessary logistics tools to F-35 program participants to facilitate efforts to operate and sustain the aircraft. ALIS consists of multiple software applications designed to support different squadron activities, such as supply chain management, maintenance, training management, and mission planning. However, we have identified numerous long-standing issues with ALIS, including that the system is not user friendly and does not provide the sustainment-related capabilities that were promised.[6] Most recently, in March 2020, we found that problems with ALIS still pose significant challenges to day-to-day F-35 operations.[7]

In response to the challenges with ALIS, we reported that DOD was replacing ALIS with a new system named the Operational Data Integrated Network (ODIN).[8] Unlike ALIS, the program is developing ODIN using an Agile development process to incrementally field capabilities more quickly. According to program officials, the program office is also developing ODIN under DOD's new policy for software development as opposed to establishing a separate acquisition program for this new effort.[9] The first step in the acquisition process under this policy is to define the requirements for ODIN in a capability needs statement, which Director, Operational Test and Evaluation (DOT&E) officials stated was finalized in September 2020. Program officials also stated that some early development work has begun. Specifically, program officials stated they are leveraging past ALIS re-design efforts to help the program field this new system as quickly as possible, with initial capability currently planned for delivery to the fleet in September 2021.[10] While ODIN development is underway, the program is still developing other documents that are

[6]GAO, *F-35 Sustainment: DOD Needs a Plan to Address Risks Related to Its Central Logistics System,* GAO-16-439 (Washington, D.C.: Apr. 14, 2016); and *F-35 Sustainment: Need for Affordable Strategy, Greater Attention to Risks, and Improved Cost Estimates,* GAO-14-778 (Washington, D.C.: Sept. 23, 2014).

[7]GAO, *Weapon System Sustainment: DOD Needs a Strategy for Re-Designing the F-35's Central Logistics System,* GAO-20-316 (Washington, D.C.: Mar. 6, 2020).

[8]GAO-20-316.

[9]DOD Instruction 5000.87 Operation of the Software Acquisition Pathway (Oct. 2, 2020).

[10]According to DOT&E officials, the current program plan is to use ALIS software on ODIN hardware.

common for new developmental efforts, such as an acquisition strategy and a cost estimate.[11]

Historical Cost Drivers and Status of F-35 Program Costs as of December 2019

DOD began development of the F-35 aircraft in 2001 without adequate knowledge of its critical technologies or a solid design, as we reported in March 2005.[12] DOD's acquisition strategy also called for high levels of concurrency between development and production—building aircraft while continuing to refine and test the designs of key components—which runs counter to GAO's leading practices for major defense acquisition programs.[13] In our prior work, we identified the F-35 program's lack of adequate knowledge and high levels of concurrency as major drivers of the program's eventual significant cost and schedule growth, among other performance shortfalls.[14]

Since 2001, DOD has significantly revised the cost and schedule goals for the program three times. DOD initiated the most recent restructuring when the program's cost for each aircraft exceeded critical thresholds established by statute—a condition known as a Nunn-McCurdy breach.[15] The restructuring process concluded when DOD established a new

[11]Section 161(a) of the William M. (Mac) Thornberry National Defense Authorization Act for Fiscal Year 2021 includes a requirement that DOD submit to Congress a strategy and implementation plan for ODIN, including an identification and assessment of goals, key risks or uncertainties, system performance metrics, and costs, among other things. The act also requires DOD to provide quarterly briefings to Congress on ODIN beginning in January 2022. Pub. L. No. 116-283 §161(a)(1), (2)(2020).

[12]GAO, Tactical Aircraft: Opportunity to Reduce Risks in the Joint Strike Fighter Program with Different Acquisition Strategy, GAO-05-271 (Washington, D.C.: Mar. 15, 2005).

[13]GAO, Joint Strike Fighter: DOD Actions Needed to Further Enhance Restructuring and Address Affordability Risks, GAO-12-437 (Washington, D.C.: June 14, 2012); and Best Practices: Capturing Design and Manufacturing Knowledge Early Improves Acquisition Outcomes, GAO-02-701 (Washington, D.C.: July 15, 2002).

[14]GAO-05-271; and GAO-12-437.

[15]Section 2433 of title 10 of the United States Code, commonly referred to as Nunn-McCurdy, requires DOD to notify Congress whenever a major defense acquisition program's unit cost experiences cost growth that exceeds certain thresholds. Significant breaches occur when the program acquisition unit cost or procurement unit cost increases by at least 15 percent over the current baseline estimate or at least 30 percent over the original estimate. For critical breaches, when these unit costs increase at least 25 percent over the current baseline estimate or at least 50 percent over the original, DOD is required to take additional steps, including conducting an in-depth review of the program. Programs with critical breaches must be terminated unless the Secretary of Defense certifies to certain facts related to the programs and takes other actions, including restructuring the programs. 10 U.S.C. § 2433a.

acquisition program baseline in March 2012 that increased the program's cost estimate by $162.7 billion and extended delivery schedules 5-6 years into the future. This March 2012 revision is the current program baseline, reflecting the cost and schedule estimates to deliver the aircraft and systems and to meet the original program requirements. As of December 2019, DOD's most recent cost estimate, total acquisition costs are $397.8 billion, as shown in Table 1.[16]

Table 1: Total Aircraft Acquisition Costs Are $2 Billion More Than 2012 Baseline Estimate

	October 2001 baseline	March 2012 baseline	Difference from 2001 to 2012	December 2019 estimate	Difference from 2012 to 2019
F-35 baseline development program costs (then-year dollars in billions)[a]					
Development	34.4	55.2	20.8	56.6	1.4
Procurement	196.6	335.7	139.1	321.2	-14.5
Military construction	2	4.8	2.8	5.2	0.4
Total program acquisition	**233**	**395.7**	**162.7**	**383.1**	**-12.6**
F-35 Block 4 modernization costs (then-year dollars in billions)[a]					
Development	NA	NA	NA	13.5	NA
Procurement	NA	NA	NA	1.3	NA
Military construction	NA	NA	NA	NA	NA
Total program acquisition	**NA**	**NA**	**NA**	**14.7**	**2.1**
Sum total of F-35 baseline and Block 4 costs	**233**	**395.7**	**162.7**	**397.8**	

Source: GAO analysis of Department of Defense data. I GAO-21-226

Note: Total aircraft acquisition costs include F-35 airframe and F135 engine costs.

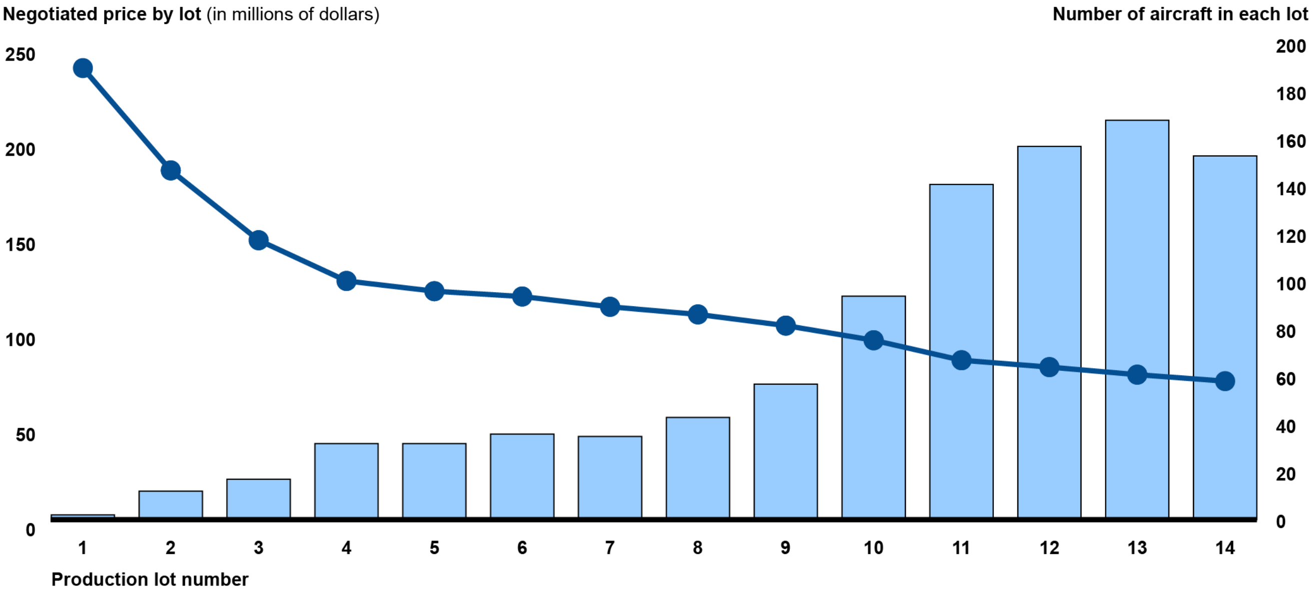
[a]Annual projected cost estimates expressed in then-year dollars reflect inflation assumptions. Baseline costs include other modernization costs that are not part of the Block 4 development program. We did not assess the reliability of the program office's F-35 baseline cost estimates. Amounts may not sum due to rounding.

From April 2019 to December 2019, the most recent published DOD estimate, the total cost estimate of the F-35 acquisition program decreased by $30 billion from over $428 billion to nearly $398 billion. This decrease was primarily due to a reduction in negotiated aircraft unit costs. In October 2019, the program finalized the contract for lots 12-14 that met

[16]We used costs from the December 2019 Selected Acquisition Report to populate table 1, the most recent costs available at the time of this report.

its goal of reducing the negotiated unit price of an F-35A to less than $80 million by lot 13, as shown in Figure 3.[17]

Figure 3: F-35A Aircraft Unit Costs Decreased Over Time



Source: GAO analysis of Department of Defense data. | GAO-21-226

Note: F-35A aircraft unit costs include the F-35A airframe and the F135 engine costs.

In May 2020, we reported that program officials stated they negotiated lower unit prices by working with the airframe contractor to leverage economic order quantity purchases and invest in cost reduction initiatives.[18] Economic order quantities involve the contractor making large purchases of components that it will use across multiple procurement lots of aircraft to reduce production costs by achieving economies of scale.

In addition, the program office and prime contractors have continued to invest in various initiatives to lower production costs. Specifically, the program office spent $320 million in efforts to improve manufacturing processes that it estimates could result in up to $10.5 billion in savings

[17]Aircraft are procured in groups, also known as production lots. Currently, DOD has negotiated contracts for aircraft through lot 14, which procures aircraft to be delivered through 2022.

[18]GAO-20-339.

over the life of the program (through 2077). Lockheed Martin received about $170 million to further lower its production costs. Pratt & Whitney estimates that $233 million in government and its own investments have yielded $7.3 billion in cumulative propulsion savings.

In addition to the acquisition costs above, the program estimates that the sustainment costs to operate and maintain the F-35 fleet for its planned 66-year life cycle are $1.2 trillion, bringing the total cost of the F-35 program to nearly $1.6 trillion.

F-35 Block 4 Modernization Effort

As we have previously reported, even though operational testing of the baseline program remains ongoing, the F-35 program office has turned some of its attention to Block 4 modernization activities and is pursuing this modernization using a different development approach.[19] DOD refers to this approach as C2D2, which is loosely based on the Agile software development process. With this approach, the program plans to deliver capabilities to the warfighter faster than it did during the baseline development program. For example, rather than take years to develop and deliver all the required capabilities to the warfighter at one time, the program intends to incrementally develop, test, and deliver small groups of capabilities every 6 months. Examples of these capabilities include a technology to avoid aircraft collisions and radar enhancements. In January 2018, to transition from the baseline development program to its Block 4 activities, the F-35 program started using the C2D2 approach to develop and test software updates to address deficiencies identified during testing.

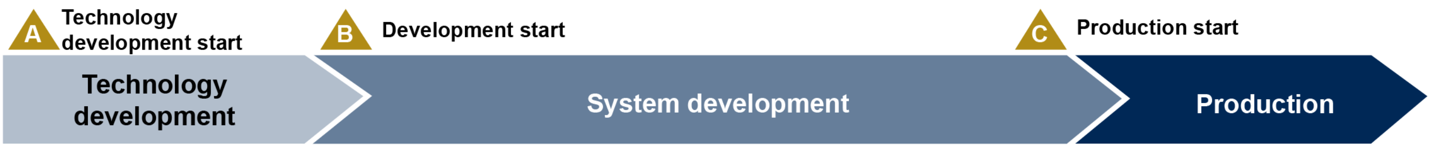
The over $14 billion cost of the Block 4 effort, by itself, exceeds the statutory and regulatory thresholds for what constitutes a major defense acquisition program, and Block 4 is more expensive than many of the other major weapon acquisitions already in DOD's portfolio.[20] To provide better oversight into Block 4 activities, in 2016, we recommended that the Secretary of Defense hold a milestone B review—a critical point in an acquisition program leading to the system development phase—and

[19] GAO-20-339.

[20] Major defense acquisition programs are those identified by DOD or that have a dollar value for all increments estimated to require eventual total expenditure for research, development, test, and evaluation of more than $480 million, or for procurement of more than $2.79 billion, in fiscal year 2014 constant dollars. See DOD Instruction 5000.02T, *Operation of the Defense Acquisition System* (Jan. 7, 2015) (incorporating change 7, April 21, 2020). See also 10 U.S.C. § 2430.

manage it as a separate major defense acquisition program, as shown in figure 4.[21] DOD did not concur with our recommendation, and it continues to manage Block 4 within the larger F-35 program. Congress subsequently passed the NDAA for Fiscal Year 2017, which it amended in 2020, that contained a requirement for DOD to submit a report containing certain elements of an acquisition program baseline—in essence, a full program business case—to include the cost, schedule, and performance information for Block 4.[22]

Figure 4: DOD Milestone Acquisition Process



Source: GAO analysis of Department of Defense (DOD) data. | GAO-21-226

In April 2019, we found that the F-35 program started Block 4 development without a complete business case identifying baseline cost and schedule estimates, which was inconsistent with leading acquisition practices.[23] Therefore, we recommended that the Secretary of Defense ensure that the F-35 program office completes its business case for the initial Block 4 capabilities under development before initiating additional development work. To date, the F-35 program completed nearly all of the documentation that is required of major defense acquisition programs; although it completed some of these documents after Block 4 development began. For example, the F-35 program office drafted, completed, or updated baseline documentation for the following acquisition documents: Acquisition Program Baseline, Acquisition Strategy, Capability Development Document, Cost Analysis Requirements Description, Independent Cost Estimate, System Engineering Plan, and Test and Evaluation Master Plan, among others.

[21]GAO-16-390.

[22]Pub. L. No. 114-328, § 224(d) and Pub. L. No. 116-92, §166.

[23]GAO-19-341.

F-35 Program Office Has Faced Testing Delays, Leading to Postponed Production Milestone

We found that F-35 simulator delays continue to prevent DOD from completing initial operational testing and making a decision to move to full-rate production. The program office postponed a full-rate production decision from the previous plan of December 2019 to March 2021, which is now further delayed to a future unknown date, and continues to take steps to address ongoing risks such as:

- high overall open deficiencies,
- production delays and quality issues,
- efforts to address Turkey's removal from the supply chain and find new suppliers, and
- aircraft not meeting reliability and maintainability goals.

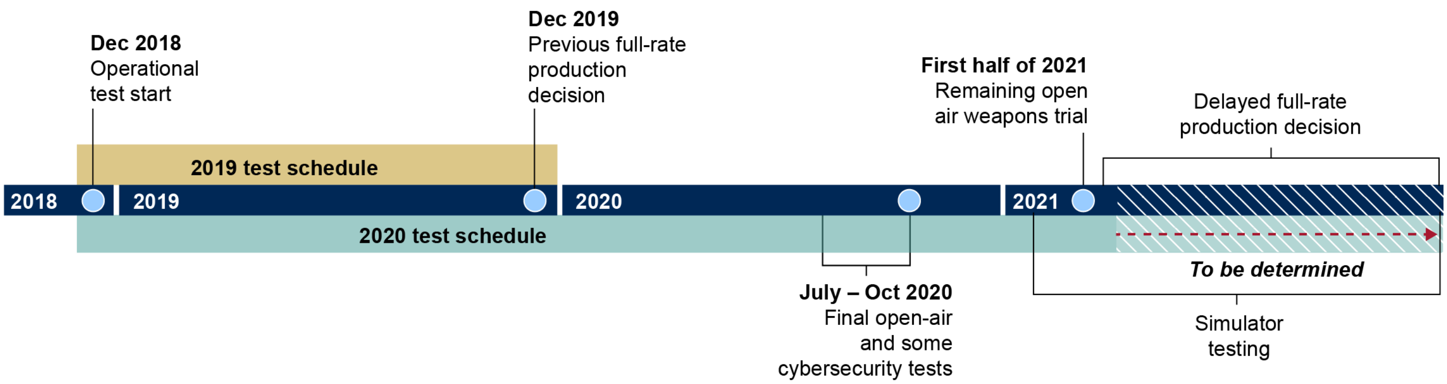
Testing Simulator Delays Drive Production Milestone Delay

We found the program did not complete its planned initial operational testing in 2020 due to delays in developing the F-35 Joint Simulation Environment, which we refer to as the aircraft simulator.[24] The simulator runs the F-35's mission systems software along with other software models (such as other weapons and modern threat systems) to provide complex test scenarios that the program cannot replicate in a real- world environment. While the 64 simulated tests required to complete operational testing will not be conducted until sometime in 2021, the program made progress in other key testing areas, including conducting three of the final four flight tests and cybersecurity testing.[25] Figure 5 shows the test schedule as of November 2020, the delay to the schedule into 2021, and the remaining test events planned.

[24] The simulator is a compilation of several aircraft, weapons, and environment effects integrated as a simulation, training and test capability.

[25] GAO-20-339.

Figure 5: F-35 Operational Test Schedules and Key Events through 2021, as of November 2020



Source: GAO analysis of Department of Defense data. | GAO-21-226

Testing officials identified technical problems with the simulator in August 2020 and have not established a time frame for fixing those problems, which has delayed its next production milestone decision. We reported in May 2020 that the program planned to complete development of the simulator and complete the simulated flights to verify and validate various threats by August 2020.[26] However, according to DOT&E officials, after conducting a readiness assessment in August 2020, the F-35 operational test team concluded that some key technical issues were not resolved. Program officials also stated that Coronavirus Disease 2019 (COVID-19) and technical complexity also contributed to delays with addressing these technical issues. DOT&E officials stated they are not considering deferring any additional testing or granting a waiver to any test requirements needed for their final report. As a result, the F-35 program office is leading the simulator's development team on an effort to create a new schedule and to identify what steps must be taken to address the technical issues and ensure that the simulator fully represents F-35 aircraft. DOT&E and program officials stated that they do not expect to develop a new schedule until spring 2021. As a result, the program will not complete operational testing in January 2021 and will not achieve a full-rate production decision in March 2021 as planned. The program has delayed those two events but has not identified a new date for completing them.

Despite these delays, the program made progress in meeting other operational testing requirements. For example, the program completed the four remaining open-air tests at Point Mugu Sea Range in California

[26] GAO-20-339.

in July 2020, which was several months later than planned. According to DOT&E officials, these delays were associated with COVID-19 restrictions and other delays with the integration of the Radar Signal Emulator threat emitter—test assets that simulate long-range threat radars—that were essential to completing those open-air tests.

The F-35 program was also able to complete two of three missile tests in 2020. The remaining missile test event is scheduled for the first half of calendar year 2021 after a planned update to the F-35 software, which DOD test officials stated would address a technical problem that has delayed the missile test event.

The program also completed the remaining initial operational cybersecurity testing on ALIS and the aircraft in October 2020. DOT&E officials said they continue to accomplish cybersecurity follow-on operational test events that are not required to complete initial operational test and evaluation, but the results will be included in the final report, if available.

Deficiencies Remain High

As of November 2020, the F-35 program had 872 open deficiencies, which is slightly higher than the 870 we reported in May 2020. Deficiencies represent specific instances where the weapon system either does not meet requirements or where the safety, suitability, or effectiveness of the weapon system could be affected. The test officials categorize deficiencies according to their potential effect on the aircraft's performance.

- Category 1 deficiencies are critical and could jeopardize safety, security, or another requirement.
- Category 2 deficiencies are those that could impede or constrain successful mission accomplishment.

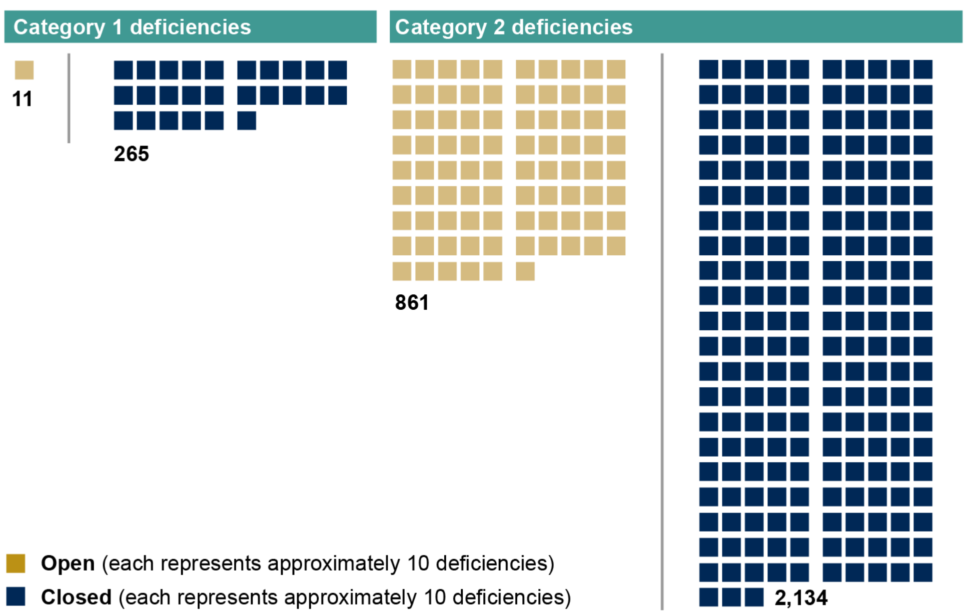
In June 2018, we recommended that the program resolve all critical deficiencies before making a full-rate production decision, in part, to reduce the potential for additional concurrency costs stemming from continuing to produce aircraft before testing is complete.[27] DOD concurred with our recommendation and stated that the resolution of

[27] GAO-18-321.

critical deficiencies identified during testing will be addressed prior to the full-rate production decision.[28]

In 2020, the F-35 program resolved 33 of the deficiencies it had identified in developmental and operational testing but it continues to find more. Of the 872 open deficiencies, the program characterizes 11 as category 1 and 861 as category 2. This represents two more open category 1 deficiencies than we reported in May 2020. According to DOT&E officials, additional new discoveries are due to quality problems with the F-35 software, resulting in a high rate of deficiency discoveries during operational testing and by pilots in the field. According to program officials, seven of these open category 1 deficiencies will be resolved prior to the completion of operational testing. Four will not be addressed until the third quarter of 2021. Figure 6 shows the total number of category 1 and 2 deficiencies that the program has opened and closed since testing began in December 2006.

Figure 6: Total Open and Closed Category 1 and 2 Deficiencies, From Start of Testing to November 2020



Source: GAO analysis of Department of Defense data. | GAO-21-226

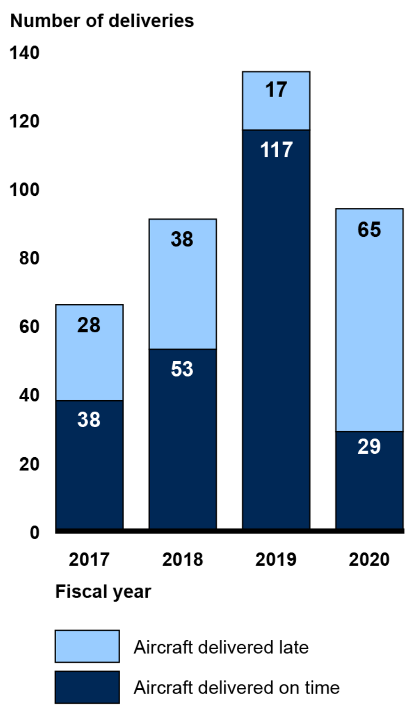
[28]GAO-18-321.

Near-Term Aircraft and Engine Quantities Reduced with More Late Deliveries

Fewer Aircraft Delivered; More were Late and Production Quality Concerns Remain

In 2020, the program reduced the quantity of aircraft that are to be delivered this year and also delivered many aircraft later than planned. The F-35 program and the airframe contractor—Lockheed Martin—planned to reduce the number of aircraft to be delivered in 2020 from 141 to 124, in part due to COVID-19-related labor disruptions and supply chain problems. As of November 2020, the airframe contractor reported delivering 94 aircraft out of 141 on contract and of those, 65 were delivered late, which was more than were late in 2019. Figure 7 shows the aircraft deliveries.

Figure 7: More Aircraft Reported Delivered Late in 2020 than in 2019



Source: GAO analysis of Lockheed Martin data. | GAO-21-226

Note: Data current as of November 2020.

Late aircraft deliveries are due to technical issues and supplier issues. DOD officials attribute these late deliveries to ongoing issues we have previously reported on, such as fastener quality problems and parts shortages, as well as emerging challenges associated with fuel tank damage and COVID-19. For more information on these and other technical risk issues, see appendix III.

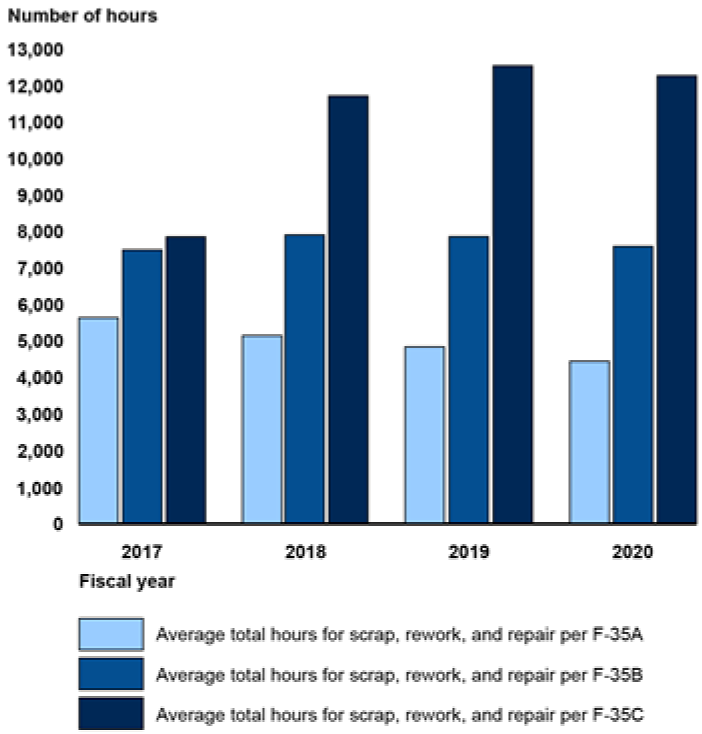
- **Titanium fasteners.** In May 2020, we reported that titanium fasteners were installed in an area of the aircraft where the design calls for a fastener stronger than titanium.[29] According to Defense Contract Management Agency officials who oversee production, manufacturing delays associated with diagnosing and fixing this problem continued to contribute to delays in 2020.
- **Fuel tank system damage.** The program found that the Onboard Inert Gas Generation System—the system in the fuel tank that replaces part of the oxygen with nitrogen to protect the tanks against explosion—was cracked on some F-35As in depots for maintenance. This cracking creates risk of the fuel tank igniting if struck by lightning during a thunderstorm. The Defense Contract Management Agency officials indicated that Lockheed Martin paused delivery of F-35 aircraft while they looked into the issue on the production line, contributing to late deliveries. In April 2020, according to DOT&E officials, the Air Force temporarily restricted F-35A from operating in weather conditions where lightning strike could possibly occur. The program developed a two-phase plan to provide relief to the F-35A lightning restriction through either inspections or another method of verifying the system's integrity.
- **Reduced capacity at suppliers**. According to Lockheed Martin and Defense Contract Management Agency officials, COVID-19 workforce restrictions at supplier facilities also led to production delays. While we have previously reported on longstanding supply chain challenges, such as late parts or parts shortages, these issues were exacerbated by COVID-19.[30] Program officials stated that Lockheed Martin conducted a supply chain assessment of impact resulting from COVID-19 and identified 37 parts challenges. According to program officials, the contractor does not expect to recover from all of these parts challenges until late 2022.

[29]GAO-20-339.

[30]GAO-20-339; and GAO-19-341.

Despite the delays to aircraft deliveries, other production metrics associated with the airframe slightly improved in 2020. Aircraft are taking less time to build, on average, for all variants, and the contractor is spending less time on scrap, rework, and repair, as shown in Figure 8.

Figure 8: Average Total Hours for Scrap, Rework, and Repair for Each F-35 Aircraft Variant



Source: GAO analysis of Department of Defense data. | GAO-21-226

Note: The program is producing and delivering three variants of the F-35 aircraft: the F-35A conventional takeoff and landing variant for the Air Force; the F-35B short takeoff and vertical landing variant for the Marine Corps; the F-35C carrier-suitable variant for the Marine Corps and the Navy.

In May 2020, we also reported that the program was not meeting all manufacturing leading practices. Specifically, we reported that 70 percent of critical manufacturing process were not in control, meaning those processes were not meeting predefined standards.[31] We recommended that the program office submit to Congress, prior to the full-rate production decision, an evaluation of the production risks associated with critical production processes that are not in control, among other things,

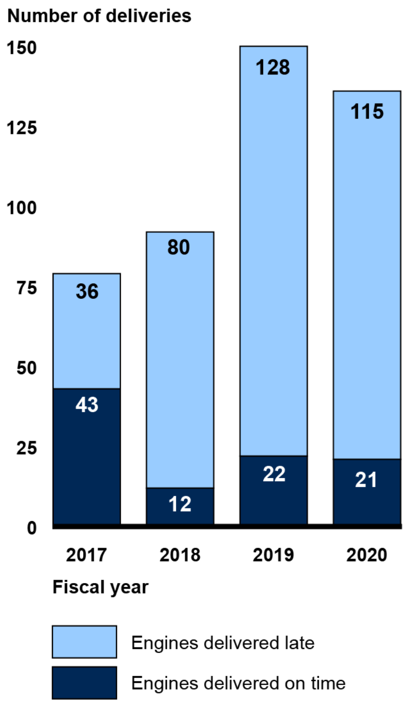
[31]GAO-20-339.

and the steps it is taking to address those risks. Since then, the F-35 program had 15 fewer critical manufacturing processes that are in control than the number we reported last year; overall, only 45 percent (3,057 out of 6,773) of processes identified are in control. According to program officials, the number of critical manufacturing processes is expected to vary over time due to the amount and types of parts produced in a given year, and a slight decrease in the quantity of manufacturing reported reflects normal variation. Officials from the Office of the Undersecretary of Defense, Acquisition and Sustainment stated they are not planning to address our recommendation because the program will keep Congress informed on these issues in quarterly updates to defense committees.

Engine Deliveries Down and Late and Quality Declined

In 2020, the engine contractor—Pratt & Whitney—continued to deliver fewer F135 engines on time, which Defense Contract Management Agency officials attribute to production quality issues and parts delays. According to Lockheed Martin representatives, even though Pratt & Whitney delivered fewer engines (21) on time as compared to the number of aircraft delivered on time (29), late delivery of these engines did not affect the aircraft delivery schedule because Pratt & Whitney builds time into its schedule to deliver the engines earlier than they are actually needed for production. As of November 2020, Pratt & Whitney had delivered 115 of 136 engines late, as shown in Figure 9.

Figure 9: F135 Engines Continue to Be Delivered Late



Source: GAO analysis of Pratt & Whitney data. | GAO-21-226

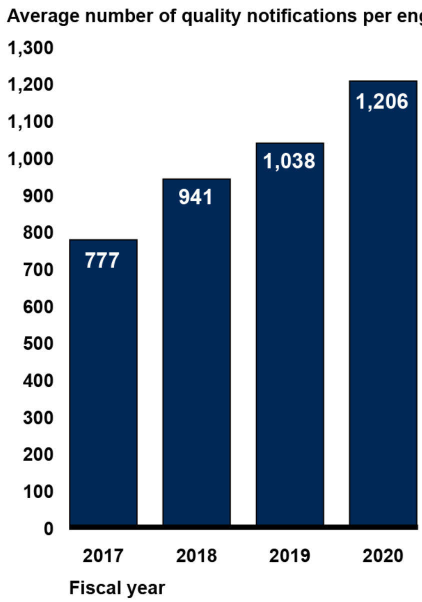
Note: Data current as of November 2020.

DOD officials stated the two main issues that affect late delivery of engines are increased demand for engine parts from fielded aircraft for flaps and seals due to coating loss and COVID-19 related effects at various suppliers.[32] For example, according to Defense Contract Management Agency officials, some suppliers responsible for critical parts of the engine were forced to briefly shut down due to COVID-19, which contributed to delays in the delivery of these critical parts to Pratt & Whitney.

As of November 2020, the average number of quality notifications per engine—production defects indicating a quality problem—was higher than in 2019, as shown in Figure 10 below.

[32]According to Pratt & Whitney representatives, the protective coating on flaps and seals—specific parts of the engine—is wearing away faster than new parts can be produced. Pratt & Whitney representatives stated they plan to increase the capacity and capability of the supplier and to implement a more durable coating—one that will last the life of the part—to help mitigate this issue.

Figure 10: Average Quality Notifications per Engine Increased in 2020



Source: GAO analysis of Pratt & Whitney data. | GAO-21-226

Note: A quality notification is an indication that a defect has been discovered.

Program Continues its Efforts to Replace Turkish Suppliers

The program continues to address supplier challenges associated with the removal of Turkey from the supply chain and has identified suppliers for 1,005 parts produced in Turkey. In July 2019, DOD removed Turkey from the F-35 program and the Under Secretary of Defense for Acquisition and Sustainment directed that the F-35 program establish alternative sources and stop placing orders with Turkish suppliers March 2020.

We reported in May 2020 that Turkey's removal from the F-35 program was likely to compound existing supply chain issues.[33] To mitigate those concerns, the Under Secretary of Defense for Acquisition and Sustainment stated the F-35 program is authorized to continue accepting delivery of parts from Turkish suppliers through the end of lot 14 deliveries (scheduled to take place through 2022).

As of December 2020, the program had identified alternative suppliers for all 1,005 parts. Furthermore, program officials stated that 95 percent of

[33] GAO-20-339.

aircraft and 76 percent of propulsion parts are qualified from the new sources and the rest are at various stages of the qualification process.[34] The program estimates it will cost $108 million to establish alternative suppliers but has not negotiated these costs with them, and therefore does not yet know what the cost impact will be for the parts being produced.

F-35 Reliability and Maintainability Is Improving but Not All Goals Are Met

We found that F-35 reliability and maintainability performance improved since our May 2020 report, but the program is still not meeting all of its performance targets. The reliability and maintainability goals lay out specific quantitative goals aimed at ensuring that an aircraft will be available for operations as opposed to out-of-service for maintenance. Last year we reported that the program was meeting 10 of its 24 reliability and maintainability goals.[35] As of June 2020, the most recent available metrics, the program was meeting or close to meeting 17 of its 24 goals. For example, the F-35A improved from not meeting its mission reliability goal, which measures the probability of successfully completing a mission of average duration, to achieving that goal at some points this year. For details about reliability and maintainability performance, see appendix IV.

Program officials attribute improvements in meeting seven more reliability and maintainability metrics in 2020 to their efforts to fund and implement reliability improvement projects over the last year. Program officials stated that they increased funding from $7 million in 2019 to $40 million in 2020 and implemented 51 new reliability and maintainability improvement projects. For example, one new project is intended to develop a new canopy coating to avoid delamination problems that program officials stated has been one of the largest drivers of lower reliability and maintainability performance. See appendix III for more details on canopy coating delamination.[36]

Although the program is still not meeting seven of its 24 reliability and maintainability goals, measurable improvements in these goals can take

[34]According to program officials, new suppliers are required to go through qualification and testing to ensure the design integrity for their parts.

[35]GAO-20-339. The program office collects 24 reliability and maintainability metrics—eight for each F-35 variant. We did not include data for one metric—the F-35C mission reliability metric—in our totals because the program did not achieve a significant sample size for F-35C mission reliability rates in 2019, according to program officials.

[36]The protective coating that covers the see-through enclosure above the cockpit, known as the canopy, was peeling (delamination).

time to manifest. For example, fielded aircraft must be modified and flown for many hours before the program can measure improvements. In our May 2020 report, we recommended that the Office of the Secretary of Defense report to the Congress on reliability and maintainability risks.[37] In response to our recommendation, DOD officials stated the F-35 program provides updates to the Defense Acquisition Executive via an Interim Program Review Defense Acquisition Board on the status of Lockheed Martin reliability and maintainability metrics.

Modernization Cost Estimates Are Increasing and Remaining Schedule, as Planned, Is Not Achievable

We found the F-35 program is now 3 years into Block 4 modernization development and the program continues to experience cost increases and schedule expansion. Costs continued to rise during 2020 due to delays in schedule and challenges in developing certain technologies, among other things. In 2020, the program added a year to its Block 4 schedule and now expects to extend Block 4 development into fiscal year 2027. We found, however, that the program office did not formulate its revised schedule based on the contractor's demonstrated past performance. Instead, the schedule is based on estimates formulated at the start of the Block 4 effort, increasing the likelihood that the scheduled 2027 completion date is not achievable.

Block 4 Development Cost Estimates Continue to Increase and Schedule Continues to Expand

The estimated cost for Block 4 development has increased and the schedule has expanded every year since the program started the development effort in 2018. Specifically, for 2020, the total reported Block 4 costs increased to reflect new development costs. In addition, the reported costs now reflect cost data for all years of Block 4 development, instead of the more limited 7-year cost estimates DOD had previously reported to Congress. For example, in 2018, DOD reported that Block 4 development would cost $10.6 billion for fiscal years 2018 through 2024. As of September 2020, DOD reported to Congress that all Block 4 costs are expected to exceed $14 billion, spanning fiscal years 2013 through 2027.

This decision to change how costs are reported stems from DOD efforts to respond to our recommendation that it improve transparency into the total costs associated with Block 4 development.[38] Specifically, in May 2020, we found that DOD's Block 4 reports to Congress, required by Section 224 of the NDAA for Fiscal Year 2017, did not fully represent the

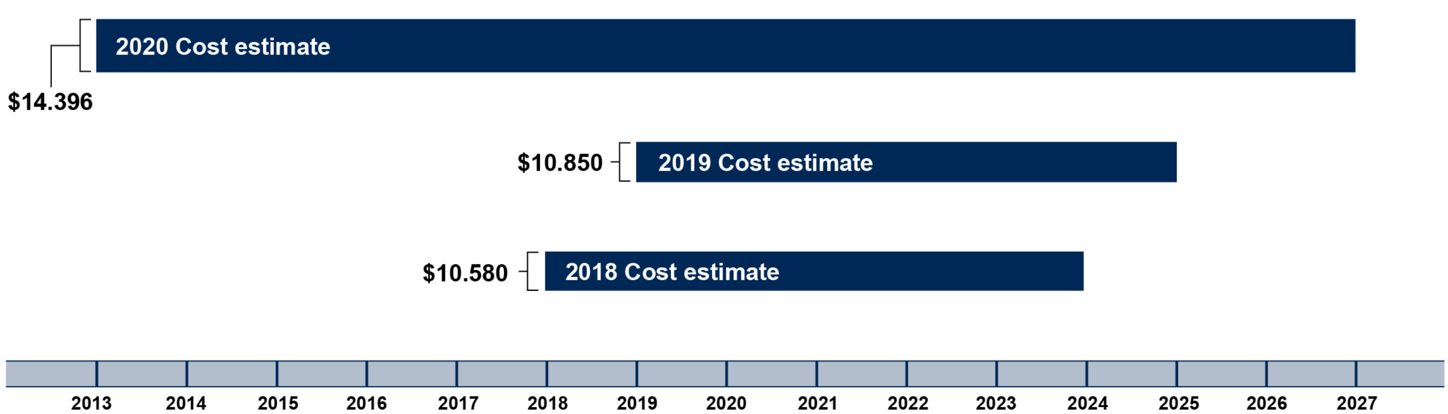
[37]GAO-20-339.

[38]GAO-20-339.

total estimated costs of Block 4 development.[39] DOD focused its reporting of Block 4 costs on the future year defense program and excluded previously incurred costs and any costs expected to be incurred after a 7-year period. For example, in May 2019, DOD reported to Congress that Block 4 would cost $10.85 billion from fiscal years 2019 through 2025. However, we identified that costs were incurred prior to fiscal year 2019 and that it had planned to continue Block 4 development through 2026.

In May 2020, we recommended that the program office provide a more holistic perspective of the total Block 4 development costs to provide Congress with improved oversight of Block 4 costs.[40] In response to our recommendation, the F-35 program began including prior year costs and future cost estimates in its Block 4 report. Therefore, in its September 2020 Block 4 report to Congress, the program's reported total cost of $14.4 billion reflects not only earlier incurred costs but also an additional 3 years of Block 4 development. Figure 11 shows the increases in both Block 4 development time frames and estimated cost.

Figure 11: Block 4 Development Cost Increased and Schedule Grew Since 2018 (dollars in billions)



Source: GAO analysis of Department of Defense data. | GAO-21-226

Note: The 2018 and 2019 estimates reflect a 7-year time frame as DOD focused its estimates on the future year's defense program—which is DOD's projected spending for the current budget year and at least the next 4 years—while the 2020 estimate includes costs for the entirety of the program, including all prior years' actual costs and the 3 additional years estimated to completion from the original 2018 estimate.

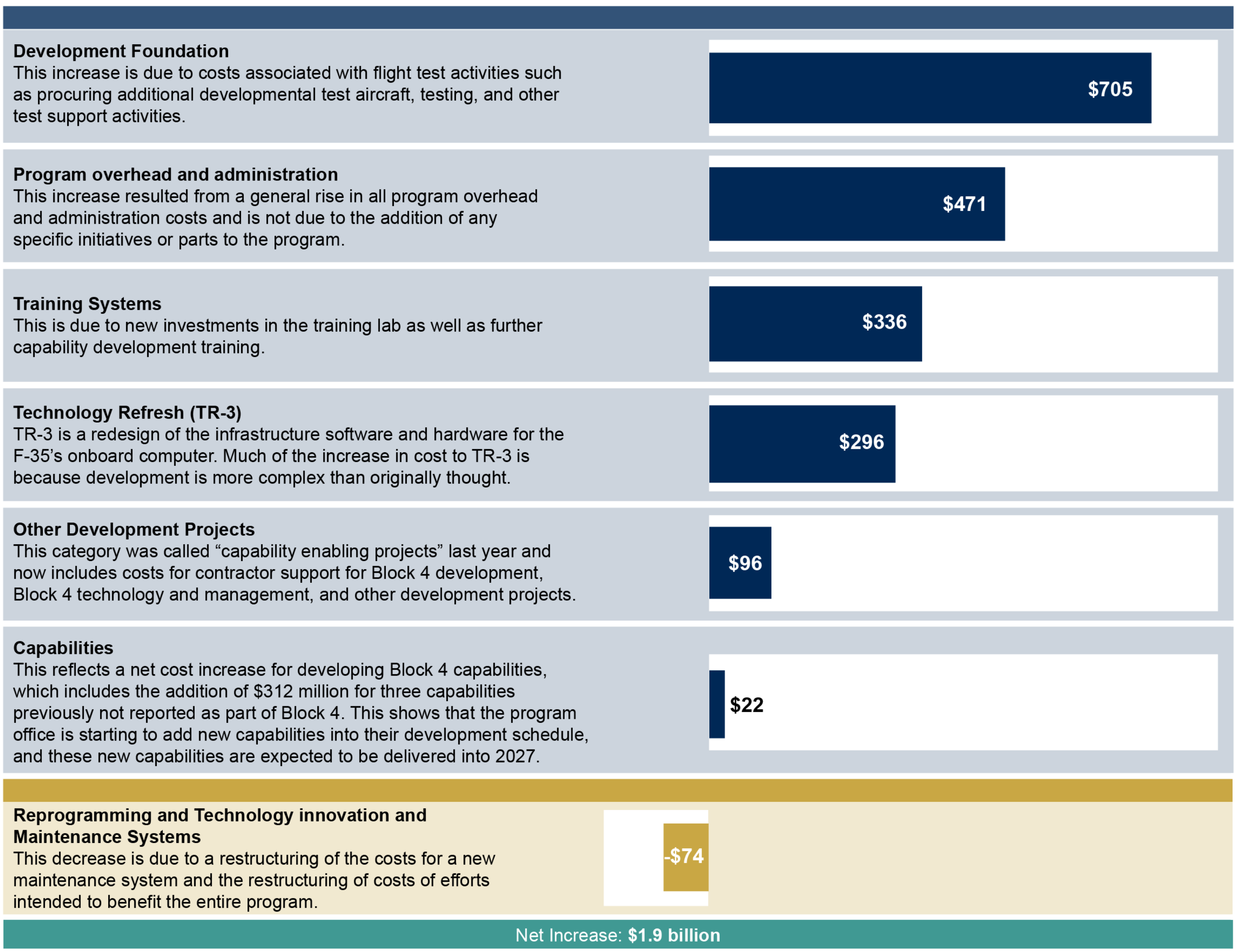
[39]National Defense Authorization Act for Fiscal Year 2017, Pub. L. No. 114-328, § 224(d), (2016).

[40]GAO-20-339.

We found that the Block 4 development cost estimate increased by $3.5 billion since DOD's May 2019 Block 4 report to Congress. Part of this cost growth is attributable to the inclusion of prior years' costs, as discussed above, and the addition of a 2-year schedule increase to complete the Block 4 program. However, over half of the increase since we reported last year—$1.9 billion—is net cost growth within various aspects of the Block 4 development program, as shown in Figure 12.[41]

[41]In May 2020, we reported that Block 4 costs were estimated to be $12.1 billion between fiscal years 2018 and 2026, based on information provided by the F-35 program. Since then, the program office identified additional years of cost, including those incurred since 2013 as well as costs through 2027. We identified that about $1.9 billion of costs in the F-35 program office's 2020 cost estimate is attributed to cost growth since we reported in May 2020.

Figure 12: Reasons for Block 4 Development Net Cost Increase by Category from May 2019 to May 2020 (dollars in millions)



Source: GAO analysis of Department of Defense data. | GAO-21-226

Key Technology Upgrade Remains on Track for 2023 Delivery

Technology Refresh 3 (TR-3)—a critical enabler of Block 4 capabilities—costs continue to grow but this effort remains on track for insertion into lot 15 aircraft in 2023. TR-3 is the suite of software and hardware technologies that will provide updated processing capability, display units, and increased memory to the aircraft. Program officials consider TR-3 to be a critical enabler to future Block 4 capabilities that are expected to be delivered starting in 2023 because those capabilities cannot function on the current hardware, known as Technology Refresh 2 (TR-2). Total costs for TR-3 increased 48 percent, or $296 million, since May 2019.[42] According to program officials, much of the increase in TR-3 costs is because its development is more complex than originally thought. While TR-3 components will still be added to lot 15 aircraft in 2023, development is currently tracking 7 months later than originally planned, according to program officials.

Program officials stated that they purchased 30 additional legacy TR-2 kits, which will mitigate the risk of late deliveries and ensure that aircraft production and testing can continue, which has also contributed to the overall cost increase. According to program officials, these TR-2 kits can be replaced with TR-3 components prior to being delivered to the government.[43] Officials also acknowledge that any further delays in TR-3 development could result in a corresponding delay to Block 4 capabilities that require the TR-3 hardware.

Other Modernization Costs Generally Increased

In addition to the $14.4 billion Block 4 costs shown in Figure 11, the program office's report to Congress identified that it will cost another $3.1 billion for the development of other aspects of F-35 modernization that are not part of Block 4. These "non-Block 4" costs had a net increase of $1.1 billion since 2019 and include increases in areas such as planning for capabilities beyond Block 4, training systems for future capabilities, and the next technology refresh, referred to as TR-4.

At the same time, a subset of these non-Block 4 modernization costs decreased over the last year. For example, program officials estimate that the costs to develop ODIN—the rearchitecture of Autonomic Logistic Information System (ALIS)—will be $5 million less than expected in 2019, although this estimate could change in the future. Specifically, according to program officials, they defined ODIN requirements in September 2020 and therefore would need to re-assess estimated ODIN costs.

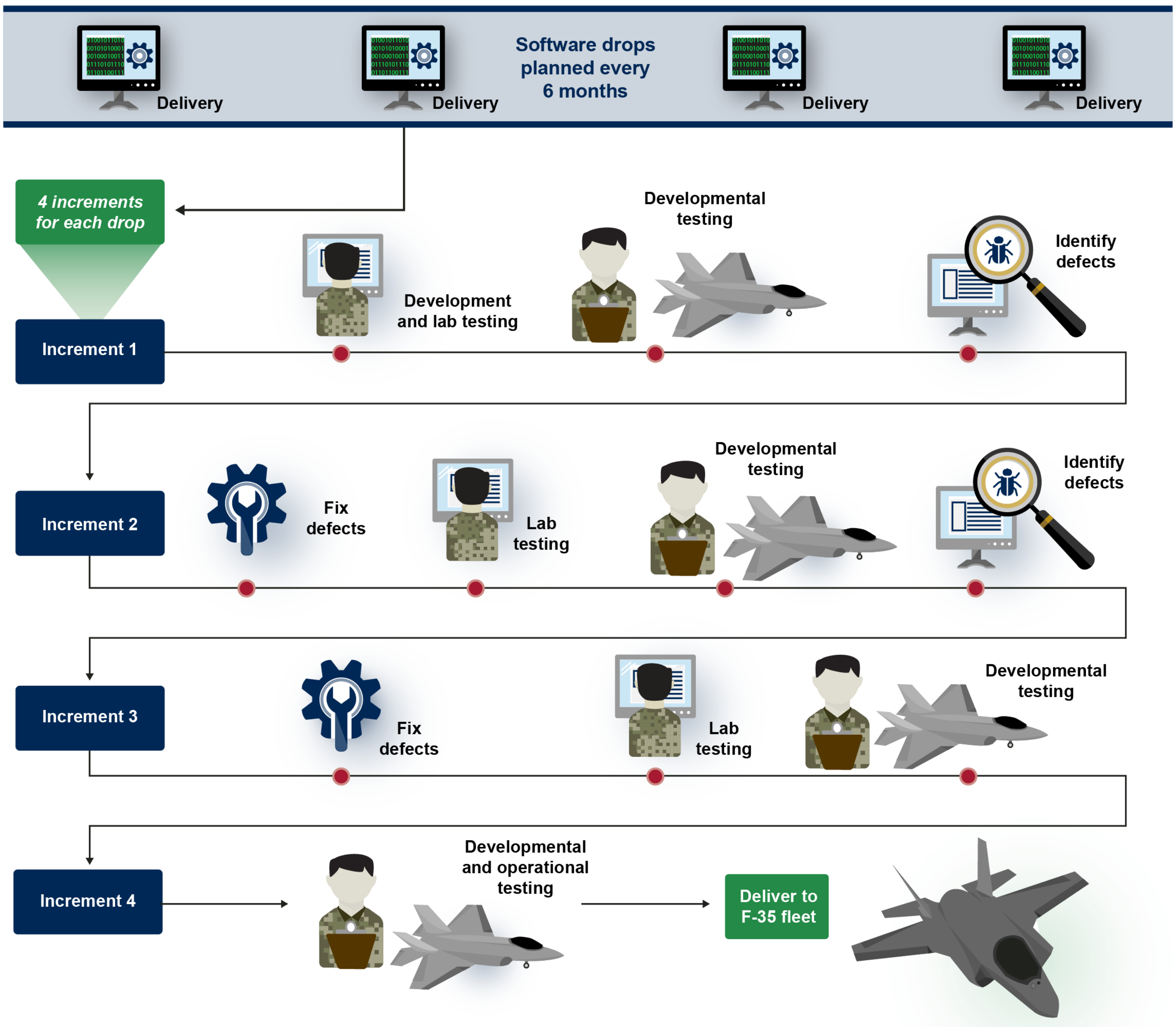
[42]The TR-3 system consists of the previous TR-2 baseline with changes to the Integrated Core Processor, Aircraft Memory System, Panoramic Cockpit Display-Electronics Unit, and Panoramic Cockpit Display-Display Unit subsystems and associated infrastructure as well as re-hosted Mission Systems applications in the new infrastructure.

[43]Currently, F-35 aircraft are using a legacy TR-2 system that is not powerful enough to enable future Block 4 capabilities, generally those that the program expects to deliver beginning in January 2023.

Block 4 Development Schedule, as Planned, Is Not Achievable

The F-35 program is more than 3 years into Block 4 development, but it has not delivered new capabilities as planned. Further, the remaining development schedule, as planned, is not based on the most recent data available and is not achievable. Under the C2D2 development approach, the F-35 program office plans to incrementally develop, test, and deliver smaller groups of capabilities to the F-35 fleet— delivered aircraft that are operating around the world—every 6 months. Figure 13 represents a notional depiction of how the C2D2 development and test process is intended to work for each 6-month software drop.

Figure 13: Notional Block 4 Iterative Development Test and Delivery Schedule



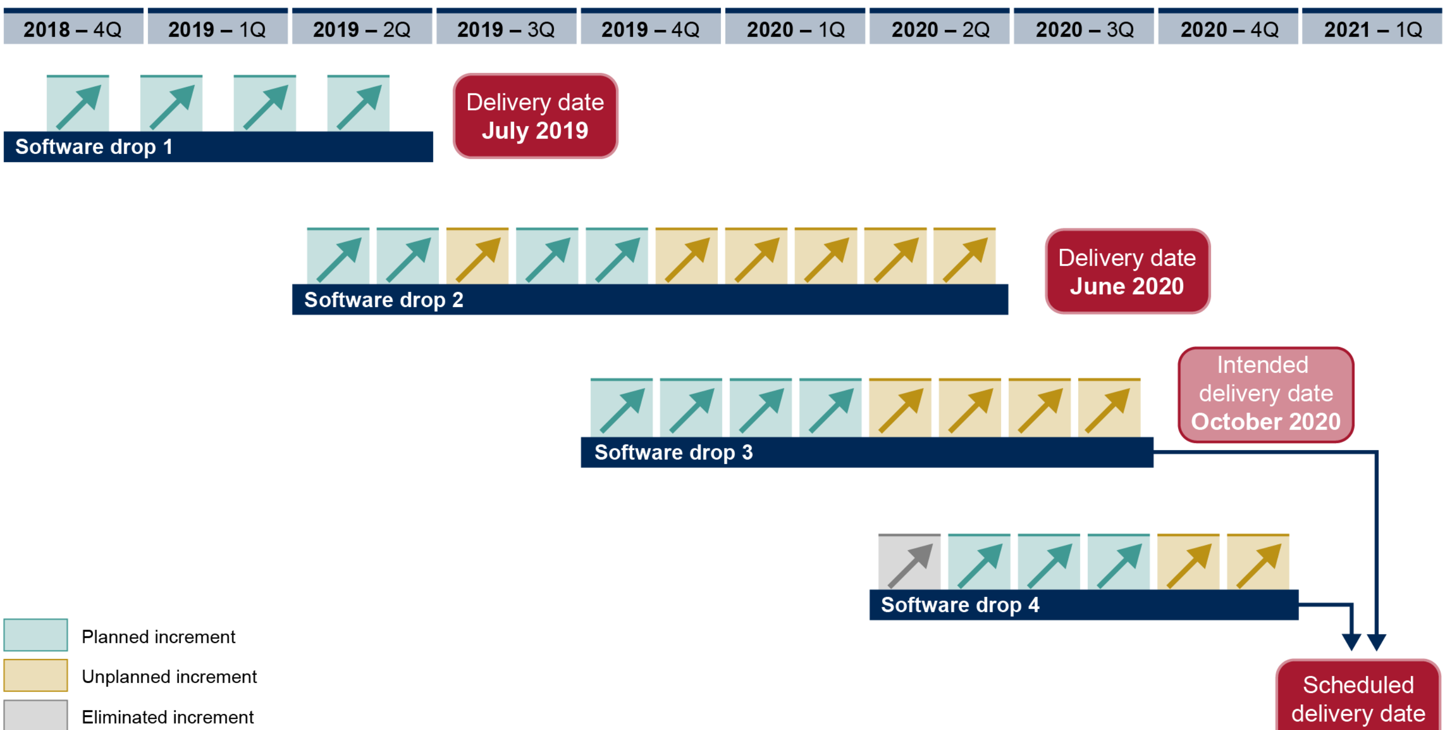
Source: GAO analysis of Department of Defense information. | GAO-21-226

As seen in Figure 13, Lockheed Martin is expected to, sequentially, develop four software increments on the way to each 6-month software drop. These increments are intended to refine and further develop capabilities over time as each is tested by the developmental test fleet. Test pilots identify potential deficiencies in each increment and report those to the contractor and the program office so that the defects can be fixed in the next increment. According to the program office, each sequential increment has a specific purpose, generally described as:

- Increment 1 should contain all new capabilities for the software drop so initial testing may proceed as planned;
- Increment 2 should address any identified deficiencies found during testing of Increment 1 and mature capabilities as needed;
- Increment 3 should address any identified deficiencies found during testing of increment 2 and mature capabilities as needed; and
- Increment 4 should be a production ready version of the 6-month software drop, capabilities should be mature, and substantial fixes should not be needed before finalizing the software for release to the F-35 fleet.

While the program generally plans for these four increments per software drop, over the last 2 years, we found that some software drops required more increments and took longer to develop than planned, as shown in Figure 14. These additional increments delayed delivery of capabilities.

Figure 14: Actual Software Increments Exceed Number of Planned Increments



Q = quarter

Source: GAO analysis of Department of Defense data. | GAO-21-226

We found that the more recent drops, in particular, have more increments beyond the planned four. For example, software delivered in June 2020 included 10 increments—six more than originally planned. Lockheed Martin representatives said that two of these increments were added to increase functionality and mature capabilities to avoid delays in the next software drop. However, according to contractor representatives, four of the added increments were to address software defects. Furthermore, the planned October 2020 software drop included eight increments—four more than planned. Lockheed Martin representatives told us that each of these added increments was to address software defects. While contractor representatives stated that they delivered all the planned capabilities, according to program officials, they decided not to field this software drop so that all defects could be resolved. Program officials also said that they will field all of those capabilities in the next scheduled software drop in March 2021, 5 months later than planned. Additionally, according to DOT&E officials, due to these delays in the October 2020 software drop, the first increment of the following software drop—planned to be delivered in April 2021—was eliminated.

For each software drop, a third-party analysis found that Lockheed Martin did not always deliver capabilities in the first increment as planned, increasing the risk that defects are included in fielded software. Including all capabilities in the first increment of a software drop provides the contractor more time to address defects before the software drops are fielded to the fleet. While Lockheed Martin representatives agree that capabilities should be delivered in the earliest increment possible, capabilities are not always ready to be included in the first increment. The contractor representatives explained that this is the case due to a number of factors, such as late contract awards preventing them from conducting new work, supply chain issues, and recent workforce capacity issues stemming from COVID-19 restrictions.

In recent years, program officials did not identify nearly a quarter of all defects until they were already delivered to test aircraft. Ideally, according to the program office, the contractor would identify defects in the software lab or before the software is fielded to the developmental test aircraft. However, a November 2020 analysis conducted by a third-party consulting firm on behalf of the program office found that between December 2017 and September 2020, 656 software defects (or 23 percent of all software defects) were identified after the software was delivered to the test aircraft.

DOT&E officials also stated that, as currently planned, the schedule does not provide adequate time to complete regression testing to identify and address defects before the final increment of the software is complete, further increasing the risk that defects are fielded to the fleet.[44] According to testing officials, finding defects after capabilities leave the contractor's laboratory is more costly and time intensive to fix. The contractor recognizes that late discoveries are a problem and is working toward identifying and fixing defects earlier in the development process. For

[44]Regression testing is re-running tests to ensure that previously developed and tested software still performs after a change. Section 162 of the William M. (Mac) Thornberry National Defense Authorization Act for Fiscal Year 2021 includes a requirement that DOD report on the status of regression testing in its quarterly briefings to Congress. This report must include an explanation of the types and methods of regression testing completed, identification of the entities that conducted regression testing, and a list of deficiencies identified during regression testing or after delivery to the fleet. Pub. L. No. 116-283, §162 (2020).

example, contractor representatives stated that they have expanded their test laboratory facilities to increase testing capacity.

Test pilots we met with explained that when the contractor did not deliver the first increment of the software drop planned to be delivered in April 2021, they lost crucial time to identify defects early in the development cycle, which contributed to the unplanned software increments and to some defects not being identified during testing. According to the program's third-party study of Block 4 software development, program officials released software to operational aircraft that included 386 software defects later found by pilots in the field. Furthermore, eight of these fielded defects adversely affected mission essential capabilities or technical cost and schedule risks without a known workaround solution. Program officials stated that they are working with the contractor to ensure that all capabilities will be included in increment 1 of each software drop going forward.

Furthermore, test pilots and DOT&E officials expressed concerns about the feasibility of the schedule going forward as the program pushes more capabilities into later software drops. Test officials said the capabilities currently being delivered are less complex than the capabilities that are planned for later in the Block 4 schedule. A review by the Undersecretary of Defense for Research and Engineering on the feasibility of the Block 4 schedule also found that the software drops planned between 2023 and 2025 are more complex and that the overall Block 4 schedule is high risk.

Despite the knowledge that the contractor is consistently adding more unplanned increments and is routinely not delivering capabilities on time or in the first increment of a software drop, we found that the program office has not adjusted its schedule to reflect these realities, because it has maintained the desire to deliver software every 6 months. According to the GAO Agile Assessment Guide, a program's schedule should realistically reflect how long each activity will take and software development teams should examine historical performance to inform future estimates.[45] Further, the GAO Schedule Assessment Guide notes the importance of an accurate baseline schedule that can be used to trace the progress made.[46] Program officials stated that while the program has revised its schedule to deliver capabilities later than initially

[45]GAO-20-590G.

[46]GAO, *Schedule Assessment Guide: Best Practices for Project Schedules,* GAO-16-89G (Washington, D.C.: December, 2015).

planned, they have not formulated a revised schedule based on the contractor's demonstrated past performance.

Furthermore, according to a July 2020 third-party analysis of the Block 4 schedule conducted by the same third-party consulting firm noted above, the airframe contractor's planned work is too optimistic and not reflective of historical performance. For example, for a recent software drop, Lockheed Martin completed only 64 percent of planned work and delivered only 69 percent of planned functionality on time. The July 2020 study identified that the contractor's performance in this regard was consistent with its historical performance, meaning the contractor is consistently delivering less than the planned amount of work that was expected to be completed. As a result, we found that the remaining schedule is not achievable as it is based on optimistic assumptions about the amount of work that can be completed and is not rooted in reality.

Program officials stated that the program is currently reviewing the feasibility of its schedule and DOT&E officials told us that the program office is considering establishing longer time frames for each software drop, such as extending them to 1 year. Simply adding time to the development cycle, however, may not fully address the program's challenges. Without a software development schedule that reflects how much work can be accomplished in each increment based on historical performance, the program office will continue to experience Block 4 development delays, and capabilities will continue to be postponed into later software drops. Delays in capability development and delivery increase the risk that capabilities will be out of date by the time they are delivered, capability development costs will be higher, and capabilities will be delivered to the fleet with deficiencies. Ultimately, this leads to warfighters waiting longer for the capabilities they need to achieve their missions. Additionally, without an updated baseline for comparing the Block 4 schedule it will be difficult for program officials, DOD decision makers, and Congress to understand program progress.

Approach for Managing Block 4 Software Development Does Not Fully Reflect Key Leading Practices

We found the F-35 program office collects data on many Block 4 software development metrics, a key practice from our Agile Assessment Guide, but the program has not met two other key practices for monitoring software development progress. Specifically, the program required Lockheed Martin to provide certain data on software development metrics through the Block 4 contract, but it does not have access to automated tools to capture real-time performance data for these metrics to help make programmatic decisions. Finally, while program officials have set some performance targets for tracked metrics, including those to try and ensure the program is meeting its 6-month software deployment schedule, they have not set targets for software quality.

F-35 Program Office Is Adding New Metrics to Measure Software Quality

The F-35 program office collects data on software metrics to monitor Block 4 software development progress, a key practice from our Agile Assessment Guide, and is taking steps to obtain additional metrics that would provide more insight into software quality.[47] For over 20 years, we have consistently emphasized the need for organizations to collect and use data about program performance to help inform and measure organization operations and results. To that end, our Agile Assessment Guide, which identifies key practices for Agile software development, states that clear, meaningful, actionable metrics provide managers information to measure program performance, and that those metrics should be tailored depending on a program's software development needs and in accordance with their established Agile framework.[48] In addition, DOD's October 2020 software acquisition instruction states that DOD programs should develop and track software development metrics, but does not define what those metrics should be and allows for programs to determine the metrics they track.[49]

We found the Block 4 contract requires the contractor to report data on metrics for software quality, performance, cost, schedule, and staffing to the F-35 program that inform software development, but these metrics provide limited insight into aspects of software development quality under the Agile software development approach.[50] The F-35 program

47GAO-20-590G.

48GAO-20-590G.

49DOD Instruction 5000.87, Operation of the Software Acquisition Pathway (Oct. 2, 2020).

50The F-35 program uses a new Agile-like development approach for its Block 4 software development effort, and relies on software development metrics collected and reported by Lockheed Martin to monitor its software development progress.

established the initial metrics the contractor is required to report on in its November 2018 contract. Program officials told us that Block 4 development activities, at that time, were focused on resolving deficiencies from the baseline program, rather than on developing new capabilities, which influenced the types of metrics the program office decided to include in the contract. The metrics currently on contract include:

- defect containment, which measures the total number of the defects that go undetected from development until flight test and the field, and
- source lines of code, which measures the size of the software.

Since the November 2018 contract award, program officials explained that, as they transitioned to developing new capabilities, they recognized the need for more information than the contractual metrics provided and took steps to collect other metrics on software development. For example, the program office worked with the contractor to obtain data on 19 metrics in addition to those required by contract, to provide further insight into the quality and performance of software development. Examples of metrics not on contract that the program receives periodically include:

- burn up software problems metric, which provides information on how much software development work remains in the scope of the project; and
- stability of software development requirements metric, which helps the program determine the extent to which changes or additions to requirements drive schedule delays or cost overruns, among other effects.

Despite these additions, program officials acknowledged that they are not collecting all the metrics they need to better understand program risks and make more informed management decisions, but are taking steps to do so. Program officials explained that they are using guidance provided by DOD and coordinating with other program offices who have used Agile software development to identify more informative Agile software development metrics. In addition, in August 2020, the program formed a joint working group, composed of program officials and Lockheed Martin representatives, to identify what other metrics the program should collect. It also formed another group, composed of program officials and contractor representatives, to identify processes for using those metrics that would better help the program ensure that the contractor is delivering software on time with fewer defects, and identify and resolve software defects earlier in the software development process.

Based on the findings of these efforts, program officials stated that the next iteration of the Block 4 contract, expected to be awarded in December 2021, will require new metrics to help achieve its goals. The program plans to require the contractor to provide, among other things:

- the number of planned features and the number of completed features for each software increment, which would provide insight into progress against the planned schedule and help ensure that all capabilities are delivered as planned in the first increment of each software drop, and
- defect density metrics for each software drop to measure the number of defects within a software increment and to measure the quality of the software being produced by the contractor.

Program officials stated that these metrics should provide better insight into on-time delivery of capabilities and software defects, two key issues hindering the program from adhering to its development schedule. These officials also stated that they intend to include other language in the December 2021 contract that would provide the program greater flexibility to request additional metrics on software development progress from the airframe contractor.

Program Office Does Not Have Access to Automated Tools to Capture Real-Time Data for Development Metrics It Tracks

We found that the F-35 program office does not have access to automated tools to capture real-time Block 4 software development data for the metrics it monitors, which is inconsistent with our key practices and DOD's new software acquisition instruction.[51] Both our Agile guide and DOD's instruction state that programs should use automated tools to the maximum extent possible to collect data on software metrics.[52] Our past work states that automated data enable managers to make data-driven decisions and provide oversight supported by relevant, real-time information.[53] In addition, using software development tools to capture and display metrics in real time helps to ensure that performance information is frequently and efficiently communicated to program managers.[54]

[51]GAO-20-590G; and DOD Instruction 5000.87, Operation of the Software Acquisition Pathway (Oct. 2, 2020).

[52]GAO-20-590G; and DOD Instruction 5000.87, Operation of the Software Acquisition Pathway (Oct. 2, 2020).

[53]GAO-20-590G.

[54]GAO-20-590G.

We found that the program's current access to software metrics data does not enable rapid or reliable assessments of the Block 4 software development effort, which hinders the program's ability to conduct contractor oversight. Program officials told us they did not include a provision in the contract to provide them with access to automated tools to collect real-time data on the airframe contractor's progress because they intentionally kept the contract language broad to maintain maximum flexibility as the program transitioned to an Agile software development approach. While Lockheed Martin uses a software interface to collect and analyze certain software development metrics data, this software interface is not accessible to F-35 program officials. As a result, program officials depend on periodic reports, generally monthly or as requested, from the contractor to make programmatic decisions.

According to a program office award fee assessment of contractor performance from June 2019 to March 2020, access to data and information remains a point of concern for F-35 program officials because the metrics data needed to make rapid decisions are not readily available. For example, the program receives data on the amount of work completed in a software iteration, but these data are available quarterly instead of provided automatically in real-time through the contractor's software interface. The data provided by the contractor are lagging, sometimes by as much as 30 days. As a result, program officials explained that they do not have insight into real-time contractor performance. Program officials stated that they are working with the contractor to gain access to its software interface, but reported slow progress on the issue. Further, program officials stated they are considering including a requirement for access to real-time data in their December 2021 contract, which is nearly a year away, and Block 4 development continues. Without automated tools to collect real-time data on software development, program officials do not have access to information that would allow them to assess progress and mitigate development and delivery risks in a timelier manner.

F-35 program officials also identified that Block 4 metrics data provided in Lockheed Martin's reports have been inconsistently reported and included errors, which has limited their insights. Our Agile Guide notes that automated tools help programs ensure the quality of the data collected is reliable.[55] In the program office assessment of airframe contractor performance between June 2019 and March 2020, the program office

[55]GAO-20-590G.

reported errors in the contractor's data, which program officials state likely stem from human error in generating performance reports. The program stated that errors and inconsistently reported data have diminished the utility of the delivered information and did not enable a reliable assessment of the status of Block 4 development. Program officials told us that they are working with the contractor to develop guidelines to ensure that data reported to the F-35 program office are reliable. They are also working to ensure all contractor employees have a uniform understanding of each metric, including the underlying data that are required, so that the program office can arrive at common findings. However, without automated tools to collect software metrics data, the program office will continue to be at risk of receiving inconsistent or incorrect data on Lockheed Martin's software development performance.

Program Office Has Not Established Performance Targets to Monitor Software Quality

We found that the F-35 program office set performance targets for software metrics associated with cost, schedule, and staffing, but had not established targets for critical metrics that assess software quality, which does not align with a key practice from our Agile Assessment Guide.[56] Our past work highlights the importance of establishing quantifiable and meaningful performance targets for software metrics to ensure that software development efforts are supporting the program's goals and making progress toward those goals.[57] Performance targets allow program managers to assess contractor performance, identify steps to improve performance, and measure progress.[58] The Statement of Work for the Block 4 contract set a target for staffing metrics related to planned workload. However, performance targets have not been set for the two critical quality issues that are affecting Block 4 schedule: on-time delivery of all capabilities in increment 1 of each software drop and defects discovered after software is delivered to the test fleet.

- As we reported above, the contractor delivered only 69 percent of the planned functionality for a recent software drop, and capabilities are not always delivered in the first software increment of the 6-month software drop schedule. By defining end state goals (targets) for functionality and then comparing the delivered functionality to the goal, program officials can see how much further the contractor needs to go, and then use that knowledge to help ensure the contractor delivers full functionality in the first increment of a software drop.

[56]GAO-20-590G.

[57]GAO-20-590G.

[58]GAO-20-590G.

Program officials could also measure progress against the target and forecast the amount of work needed to complete the drop until the increment is complete.

- Similarly, as noted above, program officials acknowledge that they discover too many defects after the contractor delivers software to the test fleet instead of being discovered in the contractor's software laboratory. By setting targets for the number of defects that the program should identify after software is delivered to the test fleet, the program could help drive the contractor to eliminate more defects in the software laboratory.

Program officials told us they did not establish targets for critical software quality metrics at the time of contract award because they were still learning which targets would be appropriate in an Agile software development environment. They stated that they are now looking to other programs, such as the F-22 program, to understand its lessons learned about which quality performance targets help to drive improved contractor performance. Program officials stated that they are considering establishing targets for software quality metrics in the December 2021 contract discussed above and are considering linking those targets to future contractor monetary incentives, but have not developed any concrete plans or placed such targets under contract. Without performance targets for critical software quality metrics, the F-35 program office is less able to assess whether the contractor has met acceptable quality performance levels and is more at risk of not meeting its Block 4 goals.

Conclusions

The F-35 is expected to serve key roles in U.S. and allied air fleets for years to come, and many updated capabilities are expected to flow from the Block 4 modernization effort. While we recognize the challenges with transitioning to Agile development, after 3 years of effort the F-35 program continues to have issues with effectively implementing the C2D2 approach to develop and deliver Block 4 capabilities. The airframe contractor continues to deliver capabilities late, and the remaining schedule contains significant risk and is not achievable based on the pace of past performance. While the program office is committed to delivering capabilities more quickly to the warfighter, the program has not delivered on its initial iterative plan. Without an achievable schedule informed by historical performance, the program is likely to continue falling short of its expectations, and the warfighter will have to wait longer for the promised capabilities.

Underlying these challenges, the F-35 program office has stated that it does not have the information on the airframe contractor's Block 4 software development performance it needs to more effectively manage the effort. While the F-35 program is taking steps toward collecting the additional metrics, only time will tell if the program office identifies the right metrics to obtain the information it needs to improve its management of Block 4 development. We will continue to monitor Block 4 software development metrics and include our observations in future reports.

Further, without requiring automated tools to access real-time contractor performance data, the program will lack timely updates on the new metrics, will lack quality program data, and will operate with old or potentially erroneous data, possibly resulting in delayed delivery to the warfighter. Finally, as the program office engages the contractor to identify the full range of needed metrics, the program has the opportunity to include software development performance targets for critical software quality metrics in the next contract to better ensure the contractor meets the program's objectives.

Recommendations for Executive Action

We are making the following three recommendations to the Department of Defense:

The Undersecretary of Defense for Acquisition and Sustainment should direct the F-35 program office to update its Block 4 schedule to reflect historical performance, to develop more achievable time frames for Block 4 modernization capability development and delivery, and to provide an accurate baseline for comparing future cost estimates. (Recommendation 1)

The Undersecretary of Defense for Acquisition and Sustainment should direct the F-35 program office to identify and implement automated tools to enable access to real-time data for software development metrics to inform program decisions and ensure the quality of data is reliable. (Recommendation 2)

The Undersecretary of Defense for Acquisition and Sustainment should direct the F-35 program office to set software performance target values for critical software quality metrics as it takes steps to identify additional software development metrics. (Recommendation 3)

Agency Comments and Our Evaluation

We provided a draft of this report to DOD for review and comment. DOD provided written comments, which we have reproduced in appendix V. In its comments, DOD concurred with all three recommendations and identified actions it was taking to address them. DOD also provided technical comments, which we incorporated as appropriate.

We are sending copies of this report to the appropriate congressional committees; the Secretary of Defense; and the Under Secretary of Defense for Acquisition and Sustainment, the Acting Secretary of the Air Force, the Acting Secretary of the Navy, and the Commandant of the Marine Corps. In addition, the report is available at no charge on the GAO website at https://www.gao.gov.

If you or your staff have any questions about this report, please contact me at (202) 512-4841 or ludwigsonj@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made key contributions to this report are listed in appendix VI.

Jon Ludwigson
Director, Contracting and National Security Acquisitions

List of Committees

The Honorable Jack Reed
Chairman
The Honorable James M. Inhofe
Ranking Member
Committee on Armed Services
United States Senate

The Honorable Jon Tester
Chairman
The Honorable Richard C. Shelby
Ranking Member
Subcommittee on Defense
Committee on Appropriations
United States Senate

The Honorable Adam Smith
Chairman
The Honorable Mike Rogers
Ranking Member
Committee on Armed Services
House of Representatives

The Honorable Betty McCollum
Chairman
The Honorable Ken Calvert
Ranking Member
Subcommittee on Defense
Committee on Appropriations
House of Representatives

Appendix I: GAO Reports and Department of Defense Actions

Table 2: Selected Prior GAO Reports on F-35 Joint Strike Fighter and Department of Defense (DOD) Responses

Year, GAO report	Estimated F-35 development costs, development length, and aircraft unit cost[a]	Key program event	Primary GAO conclusions and recommendations	DOD response and actions
2001 GAO-02-39	$34.4 billion 10 years $69 million	Start of system development and demonstration approved.	Critical technologies needed for key aircraft performance elements are not mature. We recommended that the program delay start of system development until critical technologies are matured to acceptable levels.	DOD did not concur with our recommendation. DOD did not delay the start of system development and demonstration stating technologies were at acceptable maturity levels and that it will manage risks in development.
2006 GAO-06-356	$45.7 billion 12 years $86 million	Program sets in motion plan to enter production in 2007 shortly after first flight of the non-production representative aircraft.	The program was entering production with less than 1 percent of testing complete. We recommended that the program delay investing in production until flight testing shows that the Joint Strike Fighter performs as expected.	DOD partially concurred but did not delay start of production because it believed the risk level was appropriate.
2010 GAO-10-382	$49.3 billion 15 years $112 million	The program was restructured to reflect findings from a recent independent cost team and independent manufacturing review team. As a result, development funds increased, test aircraft were added, the schedule was extended, and the early production rate decreased.	Costs and schedule delays inhibited the program's ability to meet needs on time. We recommended that the program complete a comprehensive cost estimate and assess warfighter and initial operational capability requirements. We suggested that Congress require DOD to tie annual procurement requests to demonstrated progress.	DOD continued restructuring, increasing test resources, and lowering the production rate. Independent review teams evaluated aircraft and engine manufacturing processes. Cost increases later resulted in a Nunn-McCurdy breach. Military services completed the review of capability requirements, as we recommended.
2014 GAO-14-322	$55.2 billion 18 years $135 million	The services established initial operational capabilities dates in 2013. The Marine Corps and Air Force planned to field initial operational capabilities in 2015 and 2016, respectively, and the Navy planned to field its initial capability in 2018.	Delays in developmental flight testing of the F-35's critical software may hinder delivery of the warfighting capabilities to the military services. We recommended that DOD conduct an assessment of the specific capabilities that can be delivered and those that will not likely be delivered to each of the services by their established initial operational capability dates.	DOD concurred with our recommendation. On June 22, 2015, the Under Secretary of Defense for Acquisition, Technology, and Logistics issued a Joint Strike Fighter software development report, which met the intent of GAO's recommendation.

Year, GAO report	Estimated F-35 development costs, development length, and aircraft unit cost[a]	Key program event	Primary GAO conclusions and recommendations	DOD response and actions
2016 GAO-16-390	$55.1 billion 18 years $130.6 million	DOD planned to begin what it refers to as a block buy contracting approach that was anticipated to provide cost savings. In addition, DOD planned to manage the follow-on modernization program under the current F-35 program baseline and not as its own separate major defense acquisition program.	The terms and conditions of the planned block buy and managing follow-on modernization under the current baseline could present oversight challenges for Congress. We recommended that the Secretary of Defense hold a milestone B review and manage follow-on modernization as a separate major defense acquisition program.	DOD did not concur with our recommendation. DOD viewed modernization as a continuation of the existing program and the existing oversight mechanisms, including regularly scheduled high-level acquisition reviews, would be used to manage the effort.
2017 GAO-17-351	$55.1 billion 18 years $130.6 million	The DOD F-35 program office was considering contracts for economic order quantity of 2 years' worth of aircraft parts followed by a separate annual contract for procurement of lot-12 aircraft with annual options for lot-13 and lot-14 aircraft. However, as of January 2017, contractors stated they were still negotiating the terms of this contract; therefore, the specific costs and benefits remained uncertain.	Program officials projected that the program would only need $576.2 million in fiscal year 2018 to complete baseline development. At the same time, program officials expected that more than $1.2 billion could be needed to commit to Block 4 and economic order quantity in fiscal year 2018. GAO recommended DOD use historical data to reassess the cost of completing development of Block 3F, complete Block 3F testing before soliciting contractor proposals for Block 4 development, and identify for Congress the cost and benefits associated with procuring economic order quantities of parts.	DOD did not concur with the first two recommendations and partially concurred with the third, while stating that it had finalized the details of DOD and contractor investments associated with an economic order quantity purchase and would brief Congress on the details, including costs and benefits of the finalized economic order quantity approach.

Year, GAO report	Estimated F-35 development costs, development length, and aircraft unit cost[a]	Key program event	Primary GAO conclusions and recommendations	DOD response and actions
2018 GAO-18-321	$55.5 billion 18 years $140.6 million	The program office determined that it could not resolve all open deficiencies found in developmental testing within the development program, and they would need to be resolved through post-development contract actions. DOD provided a report to Congress outlining preliminary plans to modernize the F-35. It stated it planned to develop a full acquisition program baseline for the modernization effort in 2018 and provide a report to Congress by March 2019.	The program office plans to resolve a number of critical deficiencies after full-rate production. We recommended that the F-35 program office resolve all critical deficiencies before making a full-rate production decision and identify steps needed to ensure the F-35 meets reliability and maintainability requirements before each variant reaches maturity. We also suggested that Congress consider providing in future appropriations that no funds shall be available for obligation for F-35 Block 4 until DOD provides a report setting forth its complete acquisition program baseline for the Block 4 effort to the congressional defense committees.	DOD concurred with both recommendations and identified actions that it would take in response. The National Defense Authorization Act for Fiscal Year 2019 included a provision limiting DOD from obligating or expending more than 75 percent of the appropriations authorized under the act for the F-35 continuous capability development and delivery program until 15 days after the Secretary of Defense submits to the congressional defense committees a detailed cost estimate and baseline schedule. DOD submitted its F-35 Block 4 report to Congress in May 2019, which contained cost and schedule information responding to this provision.
2019 GAO-19-341	$55.5 billion 18 years $140.6 million	For as long as the program has tracked reliability and maintainability performance, only minimal, annual improvement has been realized. Half of these metrics are failing and unlikely to meet targets outlined in the Operational Requirements Document by full aircraft maturity. As of December 2018, not all reliability and maintainability metrics within the Operational Requirements Document have been met, nor reevaluated to determine more realistic reliability and maintainability performance metrics.	We recommended that the Secretary of Defense should ensure that the F-35 program office assess the feasibility of its required reliability and maintainability targets, identify specific and measurable reliability and maintainability objectives in its improvement plan guidance, document projects that will achieve these objectives, and prioritize funding for these improvements. We also recommended that the Secretary of Defense should ensure that the F-35 program office completes its business case for the initial Block 4 capabilities under development before initiating additional development work.	DOD concurred with our four recommendations on reliability and maintainability and identified actions it would take in response. While DOD has taken some action, these recommendations are still open. DOD did not concur with our recommendation on Block 4 modernization. DOD stated that the F-35 program has adequate cost, schedule, and technical maturity knowledge to begin the development of initial Block 4 capabilities. Though these items were completed after DOD conducted additional development work, as of July 2020, the F-35 program office has completed an independent cost estimate, an approved test and evaluation master plan, and systems engineering plan. We closed the recommendation as implemented.

Year, GAO report	Estimated F-35 development costs, development length, and aircraft unit cost[a]	Key program event	Primary GAO conclusions and recommendations	DOD response and actions
2020 GAO-20-339	$57.3 billion 19 years $144.7 million	In 2019, the F-35 program conducted much of its planned operational testing but extended the schedule by 9 months, which delayed the program's full-rate production decision to between September 2020 and March 2021. In addition, the program was not meeting manufacturing leading practices identified by GAO and its Block 4 development cost estimate did not adhere to GAO leading practices.	We suggested that Congress extend DOD's Block 4 modernization reporting requirement beyond 2023 to extend to the end of the effort. We also made five recommendations to the Secretary of Defense to submit production risks to Congress prior to full rate production, to establish a Block 4 cost estimate baseline that covers all costs, and to take other steps to improve the Block 4 cost estimate. These steps are to complete a work breakdown structure, conduct a risk and uncertainty analysis, and consider technology risk assessments to help inform the Block 4 development cost estimate.	While DOD did not concur with two of our recommendations—including to evaluate production risks and update its Block 4 cost estimate with a program-level plan—it identified actions that, if implemented, will meet the intent of these recommendations. DOD concurred with our three other recommendations.

Source: GAO. | GAO-21-226

[a]The aircraft unit cost is the program's average procurement unit cost estimate, which is calculated by dividing the procurement amount by the procurement aircraft quantities. This is different than the negotiated price for F-35 aircraft, also reported above.

Appendix II: Objectives, Scope, and Methodology

This report fulfills two mandates:

- The National Defense Authorization Act for Fiscal Year 2015 included a provision for GAO to review the F-35 acquisition program annually until the program reaches full-rate production. This is the sixth report under that provision.
- The National Defense Authorization Act for Fiscal Year 2020 includes a provision for GAO to submit a report on the F-35 program's production and Block 4 progress within 30 days of the President's budget submission for Fiscal Years 2021-2025. This is the second report under that provision.

In this report, we (1) identify and describe any remaining risks with completing operational testing for the baseline program ahead of the next production milestone decision, and the steps the Department of Defense (DOD) took to mitigate those risks; (2) assess DOD's progress in developing and delivering Block 4 modernization capabilities and the program's efforts to address any remaining risks; and (3) determine the extent to which the F-35 program office is addressing selected key practices for Agile software development as it implements Block 4 development.

To identify and describe the remaining risks with the baseline program's operational testing completion ahead of the next production milestone decision and DOD's related mitigation steps, we reviewed the baseline program's costs, schedule, and performance plans and compared actual progress in each area with the goals established in its 2012 baseline to identify any significant trends. We reviewed progress on test events completed and those that remain as well as test schedules and program briefings. We conducted interviews with DOD test authorities and pilots at Edwards Air Force base and spoke with officials from the program office, the office of the Director, Operational Test and Evaluation, the officials responsible for developing the Joint Simulation Environment, Lockheed Martin (the prime airframe contractor), and Pratt & Whitney (the prime engine contractor) to discuss key aspects of operational testing progress, including flight testing results, future test plans, and progress of Joint Simulation Environment development and testing. Specifically, we obtained updates on key events required to complete operational testing.

To provide information on the program's progress with addressing remaining F-35 technical issues (identified in Appendix III) as well as with resolving deficiencies, we interviewed the same officials mentioned above and discussed changes in the number of open and closed deficiencies during 2020. We reviewed program and contractor information on deficiency reports, mitigations, resolutions, and the deficiency resolution process. To assess the program's production performance and manufacturing efficiency initiatives, we obtained and analyzed the production metrics from the program office, Lockheed Martin, Pratt & Whitney, and the Defense Contact Management Agency on their aircraft and engine delivery rates from 2017 through 2020 and discussed reasons for any delivery delays and plans for improvements.

We discussed steps taken to improve quality and on-time delivery of parts with Lockheed Martin and Pratt & Whitney representatives. We also interviewed these contractor representatives and program officials regarding the progress of identifying and validating new suppliers to manufacture parts originally produced in Turkey and associated costs. To assess reliability and maintainability of the aircraft, we obtained reliability and maintainability metrics from Lockheed Martin that were current as of June 2020, which were reported to be the most recent available. We compared the June 2020 metrics with those we included in our May 2020 report to identify any changes.[1]

To assess DOD's progress in developing and delivering Block 4 modernization capabilities and remaining risks, we reviewed program documentation, including cost and schedule estimates for Block 4 capability development and testing. Specifically, we compared the DOD F-35 Block 4 Development cost estimates for 2019 and 2020 to identify cost increases.[2] To determine the extent to which the contractor delivered Block 4 capabilities on-time and to evaluate changes to the Block 4 modernization schedule, we compared the last three revisions of the Air System Playbook (the Block 4 modernization development, test, and delivery schedule).

To determine the extent to which the program office's current Block 4 schedule is achievable, we first reviewed the GAO Agile Assessment Guide and the GAO Schedule Assessment Guide, to understand attributes that a program schedule should have. From these guides, we selected three key attributes of a high quality program schedule: (1) that a program's schedule should realistically reflect how long each activity will take, (2) that software development teams should examine historical

[1]GAO-20-339.

[2]The F-35 program office provides this annual report to Congress in response to the Section 224(d) of the National Defense Authorization Act (NDAA) for FY 2017 (P.L.114-328) Block 4 Modernization annual reporting requirement.

performance to inform future estimates, and (3) that it is important to have an accurate baseline schedule that can be used to trace the progress made. We then compared DOD's approach to developing the Block 4 schedule to determine if it met those three key attributes. We also reviewed a July 2020 Block 4 schedule analysis conducted by a third-party consulting firm that, according to program officials, was contracted by the program office to understand current schedule risk.

We interviewed officials within the program office, the Office of the Secretary of Defense Cost Assessment and Program Evaluation office, DOD test authorities at Edwards Air Force base, Defense Contractor Management Agency officials who oversee the airframe contractor, and Lockheed Martin contractor representatives to discuss the Block 4 C2D2 software development process and schedule. Specifically, we discussed the reasons why the number of planned increments for each software drop and the actual number of increments differed, the process for identifying and resolving defects associated with Block 4 software, and the progress of Block 4 capability testing and delivery. We also spoke with test pilots and Department of Operational Test and Evaluation officials about the feasibility of the remaining Block 4 schedule and reviewed the results of a report by the Undersecretary of Defense for Research and Engineering on the same topic.

To determine the extent that the F-35 program office addresses selected key practices for Agile software development, we first reviewed GAO's Agile Assessment Guide.[3] We then analyzed the Defense Innovation Board "Software is Never Done" report, the DOD's "Contracting Considerations for Agile Solutions: Key Agile Concepts and Sample Work Statement Language," and DOD's Operation of the Software Acquisition Pathway to identify similar leading practices to those highlighted in the GAO Agile Guide.[4] While there are six leading practices in the Agile Guide related to metrics, we conducted a limited assessment by focusing our analysis on the three key practices that would enable us to assess how the program uses Agile software development data to manage cost and schedule concerns that we identified in prior reports. These three key practices focus on evaluating Agile software development progress: (1)

[3]GAO-20-590G.

[4]GAO-20-590G; Defense Innovation Board, *Software is Never Done: Refactoring the Acquisition Code for Competitive Advantage* (May 2019); Acquisitions and Sustainment; DOD, *Contracting Considerations for Agile Solutions: Key Agile Concepts and Sample Work Statement Language,* Version 1.0 (November 2019); and DOD Instruction 5000.87 Operation of the Software Acquisition Pathway (Oct. 2, 2020).

tracking metrics, (2) automating real-time data collection, and (3) establishing performance targets.

To assess the extent to which the program office tracks Agile software development metrics, we obtained and analyzed F-35 Block 4 software development metrics, reviewed the Block 4 development statement of work, and reviewed briefing slides summarizing the contractor's performance in meeting contractual requirements from June 2019 to March 2020, the first performance period of the contract. We also interviewed program officials with knowledge of metrics for assessing software development, officials from the Office of the Undersecretary of Defense for Acquisition and Sustainment knowledgeable on Agile software development implementation within DOD, and Lockheed Martin contractor representatives.

To assess the extent to which the program used automated tools to collect real-time data on Agile software development, we discussed the process for obtaining data from the contractor with program office officials with knowledge of Block 4 metrics. To assess the extent to which the program office established performance targets for critical metrics, we reviewed contract documents and asked the program office to identify the performance target for each of the software development metrics provided to us and to note when no target was established.

We determined that all the data we used were sufficiently reliable for the purposes of responding to our reporting objectives. For example, we collected and analyzed the program's production data for all production lots and corroborated these metrics by interviewing contractor representatives and DOD oversight offices such as the Defense Contract Management Agency. In addition, we reviewed official program documentation on the Block 4 efforts and corroborated it through interviews with officials across DOD involved in the effort, such as the F-35 Joint Program Office cost estimating team and DOD's Cost Assessment and Program Evaluation office.

We conducted this performance audit from March 2020 to March 2021 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.

Appendix III: Status of Selected F-35 Technical Risks

The F-35 program continues to address technical risks identified in the field. Since our 2020 report, the program identified new risks with F-35A gun titanium blast panel cracking, forward engine side link bushings migration, and cracking in a system that protects the fuel tank from exploding. The program also incorporated design changes to mitigate technical risks we previously highlighted. The status of the Department of Defense’s (DOD) efforts to address these issues is as follows:

Newly Identified Technical Risks

F-35A Gun Titanium Blast Panel Cracking. Two aircraft experienced cracking in the blast panel in front of the gun. The program is conducting recurring visual inspections following each gunfire event to ensure that the cracks are not spreading and the panel is still safely in place. The F-35 program has replaced the panel with a newer panel that has larger fastener holes as an interim fix.

Forward Engine Side Link Bushings Migration. Beginning in June 2019, 17 of bushings migration were found in the engines of F-35 aircraft. Bushing migration can result in foreign object damage in the engine bay that could interfere with critical engine systems, leading to an engine shutdown as well as loss of forward engine mount structural integrity. At the time of our review, program officials did not know if the bushing migration is occurring during installation, in flight, or both. According to program officials, the program conducted a one-time inspection in April 2020. Recurring inspection criteria to mitigate the risk was being developed. Contractors modified the design to provide additional retention of the bushings as a long-term resolution.

On-Board Inert Gas Generation System Line Failure. The program found that the system in the fuel tank that replaces oxygen with nitrogen to protect the tanks against explosion was cracked on some F-35As in depot. This creates risk of the fuel tank igniting if struck by lightning during a thunderstorm. The program developed a two-phase plan to provide relief to the April 2020 F-35A lightning restriction through either inspections or another method of verifying the system’s integrity while the program office conducts a root cause investigation and develops a long-term modification.

Technical Risks Identified in Our Previous Reports

F-35B Thrust Cutback. An F-35B aircraft can experience loss of thrust during vertical landings (hover). The contractor put hover weight restrictions in place to mitigate the effect and has identified the root cause. Since September 2019, 92 percent of the aircraft have implemented all the changes to recover full capability until the very end of an engine's life. The remaining aircraft will receive the modifications required once they return to the flight line from aircraft depot modification.

F-35C Nose Landing Gear. During shipboard landings, the F-35C can experience stress that sometimes causes cracking of the coating on a part in the nose landing gear. As of June 2020, the F-35 program resolved this issue with a change that reduced excessive nose gear movement as the aircraft traveled down the launch area. The program considers this issue closed.

F-35B Three Bearing Swivel Module. The module is mounted at the back of the aircraft and allows the thrust from the engine to be redirected from straight aft for conventional flight to straight down for short take-off and vertical landing operations. In June 2019, an F-35B experienced a warning indicator in its short take-off mode due to a signal from a module. However, according to the contractor, this module should not cause a warning indicator or loss of functionality for the aircraft. The contractor has identified the root cause of the hardware issue and a gap in the software's logic that led to the warning. Both of these root causes were addressed through software that was released early in January 2020. The risk has been accepted and as aircraft are loaded, the risk is mitigated.

Canopy Coating Delamination. The contractor tested solutions for coating that was peeling in 2019 and implemented a solution of adding a vent hole to the canopy's frame. The contractor also added vent holes on production aircraft. The contractor is in the process of qualifying a second source for the canopies and full qualification is expected in the first quarter of 2022.

Helmet Mounted Display. During low-light flights, the Helmet Mounted Display's technology cannot display pure black images, instead presenting a green glow on the screen, which makes it difficult to see the full resolution of the night vision video feed. The contractor developed new display hardware to avoid this effect. According to F-35 program officials, the program placed an initial order of 62 updated displays with 35 delivered by December 2019 to support U.S. Marine Corps and Navy fleet operations. Three F-35C pilots completed initial day and night testing using the new display in July 2019 on a carrier. Program officials stated that the low-light deficiency will be addressed with a software fix in February 2021.

Appendix IV: The F-35's Reliability and Maintainability Metrics

Since 2019, the program's reliability and maintainability performance has improved for some metrics. The F-35A improved its mission reliability, mean flight hours between maintenance events, and mean flight hours between removals. The F-35B improved mean flight hours between critical failure and mean time to repair.[1] The F-35C improved mean flight hours between maintenance events, mean flight hours between removals, and mean flight hours between critical failure. The F-35C mean time to repair metric declined since 2019. The rest of the metrics stayed the same. Table 3 shows each F-35 variant's performance against these metrics' targets, as of June 2020, the most recent available metrics.

Table 3: The F-35 Reliability and Maintainability Metrics' Performance as of June 2020

	Contractually required	F-35A	F-35B	F-35C
Mission reliability[a]—measures the probability of successfully completing a mission of average duration	✓	●	●	●[b]
Mean flight hours between failure (design controlled)—measures time between failures that are directly attributable to the design of the aircraft and are considered fixable with design changes	✓	●	●	●
Mean time to repair[c]—measures the amount of time it takes a maintainer to repair a failed component or device	✓	○	●	○
Maintenance man hours per flight hour[d]—measures the average amount of time spent on scheduled and unscheduled maintenance per flight hour	✓	●	◓	●
Mean flight hours between maintenance events—also referred to as the logistics reliability metric, measures time between maintenance, unscheduled inspections, and servicing actions, including consumables[e]	—	●	◓	●
Mean flight hours between removals—measures the time between part removals from the aircraft for replacement from the supply chain	—	●	○	◓
Mean flight hours between critical failure—measures the time between failures that result in the loss of a capability to perform a mission-critical capability	—	○	●	●
Mean corrective maintenance time for critical failure—measures the amount of time it takes to correct critical failure events	—	○	○	○

Legend:
●: Metric is at or above current targets
◓: Metric is at or above minimum targets
○: Metric is below minimum targets

[1]According to program officials, in 2020 the program changed how they count critical failures. For example, previously, if an aircraft part had several components, the loss of any of those components counted as a critical failure, when in reality the loss created a degradation of the part and not a critical failure. Program officials stated that improvements for the F-35B and F-35C in mean flight hours between critical failure are attributable to this change in definition.

✓: Metric is contractually required
—: not available

Source: GAO analysis of contractor data. | GAO-21-266

Note: Each metric is measured using a 3-month average and reported on a monthly basis; this table summarizes the Joint Reliability and Maintainability Evaluation Team's review of reliability growth and maintainability improvement data from November 2009 through June 2020.

[a]Mission Reliability is a key performance parameter. Mission reliability, as well as performance against the targets related to all of these metrics, will be evaluated during initial operational test and evaluation.

[b]This is the first year we are reporting on F-35C mission reliability because, according to program officials, the number of F-35Cs that provided data is large enough to be representative of all F-35Cs in the fleet.

[c]Program officials stated that while none of the variants are at or above the current targets established in the Joint Strike Fighter Operational Requirements Document—which outlines the requirements the Department of Defense and the military services agreed the F-35 should meet—they do meet more realistic targets approved by the F-35 Joint Executive Steering Board.

[d]Maintenance man hours per flight hour is tracked as unscheduled, scheduled, and total. We report the total metric in this table because it is an F-35 Operational Requirements Document requirement.

[e]Consumable parts are nonrepairable items or repair parts that can be discarded more economically than they can be replaced or that are consumed in use (such as oil filters, screws, nuts, and bolts).

Appendix V: Comments from the Department of Defense

OFFICE OF THE ASSISTANT SECRETARY OF DEFENSE
3015 DEFENSE PENTAGON
WASHINGTON, DC 20301-3015

ACQUISITION

Mr. Jon Ludwigson
Director, Contracting and National Security Acquisitions
U.S. Government Accountability Office
441 G Street, N.W.
Washington, DC 20548

Dear Mr. Ludwigson:

The Department completed a security and accuracy review of Government Accountability Office (GAO) Draft Report, GAO-21-226, “F-35 JOINT STRIKE FIGHTER: Action Needed to Address Manufacturing and Modernization Risks,” (GAO Code 104238), and found no “FOR OFFICIAL USE ONLY” information in the DRAFT report.

The Department finds that the DRAFT report is UNCLASSIFIED and cleared for open publication, pending the GAO addressing the security and sensitivity concerns provided during the review conducted by Department’s F-35 stakeholders. Enclosed is a copy of the Department’s official security review. Additionally, comments and recommended changes to the content of the DRAFT report are enclosed.

The Department concurs with the three GAO recommendations. We have enclosed our responses.

Sincerely,

WEATHERINGTON.DYKE.DAVID.1087976218
Digitally signed by WEATHERINGTON.DYKE.DAVID.1087976218
Date: 2021.03.01 16:47:26 -05'00'

Dyke D. Weatherington
Performing the Duties of Assistant Secretary of Defense for Acquisition

Enclosures:
As stated

DRAFT GAO-21-226 (CODE 104238)

"F-35 JOINT STRIKE FIGHTER: DOD Needs to Update Modernization Schedule and Improve Data on Software Development"

Department of Defense Comments to the GAO Recommendations

RECOMMENDATION 1: The Under Secretary of Defense for Acquisition and Sustainment should direct the F-35 program office to update its Block 4 schedule to reflect historical performance and develop more achievable timeframes for Block 4 modernization capability development and delivery and to provide and accurate baseline for comparing future cost estimates.

DoD RESPONSE: Concur. The F-35 Joint Program Office (JPO) is currently assessing and updating to the Block 4 program schedule based on historical execution, budget and resourcing realities, and the insights derived from a recent independent review. We are incorporating the results of this assessment into an updated Block 4 Capability Development and Delivery schedule; expected in the second quarter CY21.

RECOMMENDATION 2: The Under Secretary of Defense for Acquisition and Sustainment should direct the F-35 program office to identify and implement automated tools to enable access to real-time data for software development metrics to inform program decisions and ensure the quality of the data is reliable.

DoD RESPONSE: Concur. The metrics to date do not fully provide a clear understanding of software maturity and quality. The F-35 JPO software team is collaborating with Lockheed Martin to identify and implement agile metrics and automated tools to better support real-time, knowledge-based program decisions. Additionally, the JPO will use these metrics and tools to assist in conducting independent assessments of F-35 software development and performance. We will establish these metrics and tools to ensure that reliable quality data is collected to better support program decisions.

RECOMMENDATION 3: The Under Secretary of Defense for Acquisition and Sustainment should direct the F-35 program office to set software performance target values for critical software quality metrics as it takes steps to identify additional software development metrics.

DoD RESPONSE: Concur. The JPO is using the results of the recent independent review, and contracting for new agile metrics, and software performance data to assess, set, and incentivize software performance targets. These efforts will be included in existing contracts where possible and in the in-work Systems Engineering Integration and Test contract.

Appendix VI: GAO Contact and Staff Acknowledgments

GAO Contact

Jon Ludwigson, (202) 512-4841 or ludwigsonj@gao.gov

Staff Acknowledgments

In addition to the contact name above, the following staff members made key contributions to this report: Justin Jaynes (Assistant Director), Jillena Roberts (Analyst-in-Charge), Jessica Berkholtz, Marvin Bonner, Rose Brister, Juaná Collymore, Erika Cubilo, Emile Ettedgui, Jennifer Leotta, Christine Pecora, Megan Setser, Roxanna Sun, Jessica Waselkow, and Alyssa Weir.

Related GAO Products

F-35 Joint Strike Fighter: Actions Needed to Address Manufacturing and Modernization Risks. GAO-20-339. Washington, D.C.: May 12, 2020.

Weapon System Sustainment: DOD Needs a Strategy for Re-Designing its Central Logistics System. GAO-20-316. Washington, D.C.: March 6, 2020.

F-35 Aircraft Sustainment: DOD Faces Challenges in Sustaining a Growing Fleet. GAO-20-234T. Washington, D.C.: November 13, 2019.

Defense Acquisitions: Observations on the F-35 and Air Force's Advanced Battle Management System. GAO-19-456T: Washington, D.C.: May 2, 2019.

F-35 Joint Strike Fighter: Action Needed to Improve Reliability and Prepare for Modernization Efforts. GAO-19-341. Washington, D.C.: April 29, 2019.

F-35 Aircraft Sustainment: DOD Needs to Address Substantial Supply Chain Challenges. GAO-19-321. Washington, D.C.: April 25, 2019.

F-35 Joint Strike Fighter: Development is Nearly Complete, but Deficiencies Found in Testing Need to Be Resolved. GAO-18-321. Washington, D.C.: June 5, 2018.

Warfighter Support: DOD Needs to Share F-35 Operational Lessons Across the Military Services. GAO-18-464R. Washington, D.C.: April 25, 2018.

F-35 Aircraft Sustainment: DOD Needs to Address Challenges Affecting Readiness and Cost Transparency. GAO-18-75. Washington, D.C.: October 26, 2017.

F-35 Joint Strike Fighter: DOD's Proposed Follow-on Modernization Acquisition Strategy Reflects an Incremental Approach Although Plans Are Not Yet Finalized. GAO-17-690R. Washington, D.C.: August 8, 2017.

F-35 Joint Strike Fighter: DOD Needs to Complete Developmental Testing Before Making Significant New Investments. GAO-17-351. Washington, D.C.: April 24, 2017.

F-35 Joint Strike Fighter: Continued Oversight Needed as Program Plans to Begin Development of New Capabilities. GAO-16-390. Washington, D.C.: April 14, 2016.

F-35 Sustainment: DOD Needs a Plan to Address Risks Related to Its Central Logistics System. GAO-16-439. Washington, D.C.: April 14, 2016.

F-35 Joint Strike Fighter: Preliminary Observations on Program Progress. GAO-16-489T. Washington, D.C.: March 23, 2016.

F-35 Joint Strike Fighter: Assessment Needed to Address Affordability Challenges. GAO-15-364. Washington, D.C.: April 14, 2015.

F-35 Sustainment: Need for Affordable Strategy, Greater Attention to Risks, and Improved Cost Estimates. GAO-14-778. Washington, D.C.: September 23, 2014.

F-35 Joint Strike Fighter: Slower Than Expected Progress in Software Testing May Limit Initial Warfighting Capabilities. GAO-14-468T. Washington, D.C.: March 26, 2014.

F-35 Joint Strike Fighter: Problems Completing Software Testing May Hinder Delivery of Expected Warfighting Capabilities. GAO-14-322. Washington, D.C.: March 24, 2014.

F-35 Joint Strike Fighter: Restructuring Has Improved the Program, but Affordability Challenges and Other Risks Remain. GAO-13-690T. Washington, D.C.: June 19, 2013.

F-35 Joint Strike Fighter: Current Outlook Is Improved, but Long-Term Affordability Is a Major Concern. GAO-13-309. Washington, D.C.: March 11, 2013.

Fighter Aircraft: Better Cost Estimates Needed for Extending the Service Life of Selected F-16s and F/A-18s. GAO-13-51. Washington, D.C.: November 15, 2012.

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 our website. Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. You can also subscribe to GAO's email updates to receive notification of newly posted products.

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, https://www.gao.gov/ordering.htm.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on Facebook, Flickr, Twitter, and YouTube.
Subscribe to our RSS Feeds or Email Updates. Listen to our Podcasts.
Visit GAO on the web at https://www.gao.gov.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact FraudNet:

Website: https://www.gao.gov/fraudnet/fraudnet.htm

Automated answering system: (800) 424-5454 or (202) 512-7700

Congressional Relations

Orice Williams Brown, Managing Director, WilliamsO@gao.gov, (202) 512-4400, U.S. Government Accountability Office, 441 G Street NW, Room 7125, Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, DC 20548

Strategic Planning and External Liaison

Stephen J. Sanford, Acting Managing Director, spel@gao.gov, (202) 512-4707
U.S. Government Accountability Office, 441 G Street NW, Room 7814, Washington, DC 20548