Federal Aviation Administration

DOT/FAA/AM-08/1
Office of Aerospace Medicine
Washington, DC 20591

Development of an Aeromedical Scientific Information System for Aviation Safety

Connie L. Peterman
Paul B. Rogers
Stephen J.H. Véronneau
James E. Whinnery

Civil Aerospace Medical Institute
Federal Aviation Administration
Oklahoma City, OK 73125

January 2008

Final Report

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-08/1		

4. Title and Subtitle	5. Report Date
Development of an Aeromedical Scientific Information System for Aviation Safety	January 2008
	6. Performing Organization Code
7. Author(s) Peterman CL, Rogers PB, Véronneau SJH, Whinnery JE	8. Performing Organization Report No.
9. Performing Organization Name and Address FAA Civil Aerospace Medical Institute P.O. Box 25082 Oklahoma City, OK 73125	10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency name and Address Office of Aerospace Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591	13. Type of Report and Period Covered 14. Sponsoring Agency Code

15. Supplemental Notes

16. Abstract

Introduction. The Bioinformatics Research Team at the Civil Aerospace Medical Institute (CAMI) uses data analysis techniques to study issues associated with medical certification decisions and their effects on the U.S. pilot population to ensure safety of flight. We developed a Scientific Information System to assist in research efforts associated with statistical and epidemiological studies of the U.S. civil pilot population. Significant data challenges exist relative to the integration and analysis of very large datasets associated with civil aviation.

Methods. The CAMI aviation safety/medical certification data warehouse was created with data from varying time periods. Data includes NTSB mishap data from 1983 to 2005, FAA Accident Incident data from 1971 to 2005, airmen registry data (combined with medical certification data) from 1962 to 2005, toxicology data from 1990 to 2005, and autopsy data from 1980 to 2005. The research methodology, developed using records from the CAMI warehouse, was used to create the Aerospace Medical Research Scientific Information System that contains new metrics for comparing groups of aviators. This was done by developing a methodology that combined the various data sources into a single integrated database while transforming the data into a format conducive to epidemiological studies.

Discussion. We will discuss the methodologies developed to create new metrics—Active Airmen, Months Contributed, and Effective Class—which show promise in comparing groups of aviators with various pathologic conditions. The distributions and evolution of pathologic conditions can be observed in the resulting Scientific Information System pilot population for the time period of interest. The Scientific Information System overcomes the data incongruities between the source databases and makes analysis possible with statistical programs.

Conclusion. CAMI was successful in creating a Scientific Information System, which is a permanent database for use in epidemiological aviation research, by integrating multiple datasets and allowing the investigation of potential safety-related isssues. The Scientific Information System was created to improve data handling issues and bring cutting edge analytical tools to allow explorations of rare outcomes and to develop risk management models. The Scientific Information System permits aviation safety-related epidemiological research on the entire U.S. civil pilot population.

17. Key Words	18. Distribution Statement
Scientific Information System, Data Mining, Decision Support System, Data Warehouse, Database, Bioinformatics, Data Analysis, Aviation Safety, Epidemiology, Aviation Medicine	Document is available to the public through the Defense Technical Information Center, Ft. Belvior, VA 22060; and the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	21	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Development of an Aeromedical Scientific Information System for Aviation Safety

We're drowning in information and starving for knowledge…
—Rutherford D. Rogers

INTRODUCTION

A Scientific Information System is defined as a computerized Management Information System (MIS) that takes a systems viewpoint and is based on scientific principles of theory construction, data quality, and systematic data analysis (Holt, 2001). The Bioinformatics Research Team created a Scientific Information System (SIS) to deal with the increasingly large government observational datasets on aviation mishaps (incidents and accidents) and airman pilot and medical certifications while incorporating system safety principles. This paper will describe the creation of the Scientific Information System which is a permanent epidemiological database designed exclusively for aviation research.

A knowledge discovery process was developed to consolidate different aviation data sources into a single dataset with a format more conducive to statistical analysis. This process involved selection, preprocessing, transformation, data mining, and evaluation of many different data sources (Dunham, 2003). The result of this work was a single system that accesses not one but many different information resources, combining them into a consolidated and integrated whole. The data in the resulting SIS represent the entire population of U.S. pilots rather than just a sample; thus, the statistical results are population parameters rather than statistical estimates and are not subject to sampling error (Holt, 2001).

The ultimate goal of a SIS is to be able to explain, predict, and precisely control the processes and outcomes of complex systems (Holt, 2001). One benefit of our SIS is that it will support epidemiological researchers in aviation safety studies who are not familiar with the underlying process of the dataflow, collection, and storage. This system will support studies undertaken that examine the aviation safety and aeromedical aspects of certifying pilots with various pathological conditions. Finding patterns in the distribution of various pathologies in the mining of the electronic exam records of the U.S. pilot population is essential in any aviation epidemiological study. From this, predictive models can be constructed.

Historical Efforts

There have been efforts at creating similar systems in the past. One such attempt, in the 1980s, involved integration efforts within the FAA using mainframe computer technology. The Medical Accident System, managed by the Air Medical Statistical Section of the Aeromedical Certification Division, sought to bring together FAA Accident and Incident data along with medical certification data. It was discontinued in the late 1980s.

Successful attempts at creating a research aviation system outside of the FAA go back as early as 1981 with the construction of an information collection system for medical data generated in U.S. Air Force acceleration trials (Whinnery & Slaughter, 1981). The resulting datasets were used to study the medical, epidemiological, and physiological responses of aviators. Systems such as these aspired to the same goals as a Scientific Information System described by Holt (Holt, 2001). Whinnery, a coauthor of the current report, was a chief architect of the U.S. Air Force acceleration trials data collection effort.

The recent history of such data integration efforts at CAMI dates back to 1990 with the initiation of the Age 60 Project. That project was a three part approach to provide the FAA with data to use in evaluating the Age 60 Rule of 1959 ,Title14 of the Code of Federal Regulations (CFR), Chapter 1, Part 121, §121.383(c)[1], which mandates the retirement of airline transport pilots on or before their 60th birthday. The three parts involved were: a review of the literature, the creation of a consolidated database in which to study the aviation epidemiology of pilots, and the planning of a longitudinal study to discover what factors regarding aging would be found among older pilots.

The original legacy aviation safety datasets that have been combined into the current SIS and previous consolidation attempts were not created with research methodology in mind. The consolidated database created in the early 1990s showed that the datasets could be linked and identified some of the issues affecting the quality of the records as well as the degree of matching attainable. It also highlighted an early data anomaly that was unknown to anyone at the time: that a significant proportion of medical records were permanently missing from 1986 (Kay et al., 1994). The consolidated database

[1]http://ecfr.gpoaccess.gov/cgi/t/text/text-idx?c=ecfr&tpl=/ecfrbrowse/Title14/14tab_02.tpl

was the precursor to the CAMI warehouse research, and one of the coauthors, Véronneau, was a consultant to the project.

The creation of the CAMI Decision Support System (DSS) warehouse in 2003 benefited from the prior data consolidation effort and incorporated improvements in computer hardware, database, and data warehousing technologies available almost a decade later. The CAMI DSS contained a subset of data, the Aviation Accident and Medical Database-Decision Support System (AAMD-DSS), which represented the core of the research interest on our part.

At CAMI, the CAMI-DSS was created initially somewhat at cross purposes between the need for an enterprise-wide (Office of Aerospace Medicine) data repository which can be easily queried by managers and a comprehensive archive of all detailed information from disparate government data sources for research users. In the end, the compromise solution fit neither group's needs completely. However, it did achieve many of the overall design objectives and established a unique data collection of medical and aviation safety information. The system should still be considered a sophisticated and stable prototype; however, the resulting complex dataset can be a challenge to master and can easily be misinterpreted by novice query attempts at relating information. The main difficulty is that the CAMI DSS is a compromise solution which resulted in a combination of numerous tables linked together with no practical way for the average user to navigate between them. The complexity of this schema limits the knowledge which can be extracted from the data. As researchers, we developed methods to organize the available warehouse data into information sets to suit particular research projects. This necessitated some complex data rollups to be exported as large rectangular datasets that are easily imported into data analysis programs. After data cleansing and quality assurance, certain additional derived variables were calculated, and the dataset was made available to researchers. The derived variables contribute much of the value of the SIS.

METHODS

A multidisciplinary team, with skill sets in aircraft piloting, accident investigation, aerospace medicine, programming, database administration, statistics, mathematics, engineering, computer hardware, software, and networking, was assembled to glean the most knowledge from the complicated aeromedical datasets. The following process was followed and serves as the basis for this initial report that documents our methodology. Further papers will follow with the results of the epidemiological studies and with the models developed.

Sources of Data

In 1999, the Document Imaging Workflow System (DIWS) became operational at CAMI. Within it are the electronic medical records of some 3 million pilots and 12 million medical certification physical exams. An FAA designated Aviation Medical Examiner (AME) performs the required examination, governed by Title14 of the Code of Federal Regulations , Chapter 1, Part 67, using the medical certificate application FAA Form 8500-8. The AME then transmits the results via the Internet to the DIWS. Some test results and examinations are mailed to CAMI for scanning and processing into the DIWS. This transactional processing and workflow information management system is based upon a relational database. Approximately 2,000 medical certification applications are received each day.

The AAMD-DSS warehouse was established at CAMI to consolidate various sources of aviation safety and aeromedical certification data to permit studies of such matched and related data. Included in the warehouse are clones of DIWS, the National Transportation Safety Board (NTSB) aviation database, the FAA Accident and Incident Data System (AIDS), the Airmen Registry pilot certificate component, and several specialized aviation safety databases developed at CAMI in the Aerospace Medical Research Division. Information from these separate systems is rolled up and linked together in the resulting Scientific Information System.

Defining Members of the U.S. Civil Pilot Population

A pilot must have an airman pilot certificate appropriate for the aircraft and flight conditions to be undertaken and must also have a valid airman medical certificate. Generally, once earned, unless suspended, revoked or surrendered, airman pilot certificates do not have an expiration date; however, airman medical certificates always have an expiration date and must be periodically renewed. (Title 14 of the Code of Federal Regulations, Chapter 1.)

Part 1.1 defines a medical certificate as "acceptable evidence of physical fitness on a form prescribed by the administrator." Any person who meets the medical standards prescribed in Title 14 of the Code of Federal Regulations (14 CFR), Chapter 1, Part 67, "based on medical examination and evaluation of the person's history and condition, is entitled to an appropriate medical certificate." Airman medical certificate examinations must be performed by an AME of which there are over 5,000[2] worldwide. An AME is authorized to receive applications,

[2]www.faa.gov/licenses_certificates/medical_certification/faq/response11/

to perform physical examinations and to issue airman medical certificates. They may also defer an application for further review or deny an applicant.

Certain medical conditions are disqualifying; however, in many cases when the condition is adequately controlled, the FAA will issue an airman medical certificate contingent on periodic reports through a special issuance process. There are no minimum or maximum ages for obtaining an airman medical certificate. Any applicant who is able to pass the exam may be issued a medical certificate. However, since 16 years is the minimum age for a student pilot certificate, people under 16 are unlikely to have practical use for an airman medical certificate.

Medical certificates have varying durations:

- A first-class airman medical certificate is required to exercise the privileges of an airline transport pilot certificate. A first-class airman medical certificate is valid for the remainder of the month of issuance plus 6 months for activities requiring a first-class medical certificate. After that time, the certificate is valid for 6 months for activities requiring a second-class medical certificate plus an additional 12 to 24 months for activities requiring a third-class medical certificate depending on the age of the pilot at the time of the exam.
- A second-class airman medical certificate is required for commercial, non-airline duties (e.g., for crop dusters, corporate pilots) and is valid for the remainder of the month of issuance plus 1 year. Those exercising the privileges of a flight engineer certificate, a flight navigator certificate, or acting as an air traffic control tower operator must hold a second-class airman medical certificate. After that time, the certificate is valid for an additional 12 to 24 months for activities requiring a third-class medical certificate, depending on the age of the pilot at the time of the exam.
- A third-class airman medical certificate is required to exercise the privileges of a private pilot certificate, recreational pilot certificate, a flight instructor certificate, or a student pilot certificate. A third-class airman medical certificate is valid for the remainder of the month of issuance plus 3 years for pilots under age 40 or 2 years for those pilots age 40 and over. Prior to the September 16, 1996 revision of Title 14 of the Code of Federal Regulations (CFR), Chapter 1, Part 61, §61.23(d), the duration of validity of the Class 3 medical certificate was two years regardless of age.

Special Issuance airman medical certificates are issued to airmen with a disqualifying medical condition who, upon further examination, are allowed to fly under special circumstances. Special Issuances often have a stipulated expiration date in the authorization letter sent to the airman by the AME, which typically decreases the period of validity for the medical certificate. The expiration date is entered in the record keeping system (DIWS) by the AME's transmission of such a date in the limitations field of the airman medical certificate and the actual date in the expiry date field by the Aerospace Medical Certification Division (AMCD), the Regional Flight Surgeon (RFS), or Office of Aerospace Medicine Headquarters. The medical records only have information regarding time limited medical certificates, such as is commonly done with initial special issuance medical certificates, from 1999 on. The use of this date in the electronic medical record is not available for records prior to 1999.

Pilots report their occupation and type of airman certificates held on the FAA Form 8500-8, however we used only airmen certificate information from the Airmen Registry dataset, which is merged and matched to the DIWS dataset in the AAMD-DSS. Only records with an airman certificate type of pilot and an airman certificate level of Airline Transport Pilot, Commercial Pilot, Private Pilot, Recreational Pilot, or Student Pilot were used. By employing the Airmen Registry data we were able to rely on a verified source of airman certificate information.

Process of SIS Construction Using Active Airman Algorithm

For an 11-year period from 1993-2003, the electronic medical records of pilots were analyzed and algorithms developed to allow us to study the prevalence of pathologies and overall counts for the U.S. civil pilot population and facilitate the determination of the number of months for a pilot's active status by year. The SQL scripts and the order of their execution in building and defining the population within the SIS are referred to as the *Active Airman Algorithm.* This algorithm operates on the CFR rules in determining who is eligible to be a member of the U.S. civil pilot population and their length of eligibility at the different medical classes. Each member's AMCD assigned pathology code, which records health information and AMCD administrative actions at the time of their medical, is recorded along with other matched pertinent data for each individual from the other government legacy data sources, such as the NTSB dataset. The data sets produced by the algorithm are more easily processed and transformed in order to overcome the data incongruities between the source data of various aviation systems. Mathematical transformation of these data sets is required when the raw data have strong asymmetry, many outliers, or large and systematic residuals. Transformation lessens some of these issues and makes analysis of the data more feasible (Hoaglin, Mosteller, & Tukey, 2000).

The process flow, depicted in Appendix A, is just a logic chain to determine the order of execution of the numerous SQL scripts developed specifically for this method. The scripts were run either from a Microsoft SQL Server front end to pull data from the CAMI warehouse or directly on the warehouse using Quest's Oracle development tool TOAD. They document many of the data manipulations necessary to achieve the formats used in the statistical analysis and data mining. The data was then imported into the SIS through an extraction, transaction and loading method.

The advantages of our Active Airman Algorithm include an accurate determination of a pilot's active status for each year and an objective determination of the number of months in that year that the pilot was active based upon examination of each electronic medical record for that pilot. This method enabled an accounting for the inflow and outflow of pilots as well as tracking pilots who developed a medical condition of interest over the study period. In the creation of the Active Airman Algorithm, the original data set was parsed from 1991 as the origin in order to correctly capture all the active pilots for the 1993 calendar year that begins this study. Exam records were also needed from 1990 to be able to calculate the expiration date of a valid medical certificate in 1992, which itself was needed to be able to calculate months contributed for 1993, the beginning of the study.

Using the class of medical issued, a two-digit code from Code Schedule A in the AMCD handbook, class was filtered to class 1 (11-19), class 2 (21-29), and class 3 (31-39) certificates that correspond to first, second, and third class medical certificates, respectively. This method of tabulating pilots with valid medical certificates excluded those who were in a pending (40) or denied (90-98) status for that year.

The resulting linked SIS database was housed in a Microsoft SQL Server 2000 database engine. The queries used to retrieve and fuse the data were written in P/L SQL for Oracle version 8 and Microsoft Transact SQL. All data analysis and modeling was performed in SAS version 9, SPlus Enterprise Developer 8, and Insightful Miner version 7. Process flow charts were done with SmartDraw Suite Edition version 7.

Static and Dynamic Variables

The Active Airman Algorithm creates placeholder records in subsequent years for airmen based on their most recent exam, allowing us to study the frequencies and distribution of pathologies throughout the U.S. pilot population. Because a pilot can appear as an eligible member of our defined population for a maximum of three years (post 1996) based on a single original medical exam, the airman's original record is repeated as a "placeholder" record. Many of the variables of interest remain the same as they did at the time of the actual exam; these are static variables. Algorithms were developed within the SIS, many of which are described in Appendix A (Figures A.1 thru A.8), that create dynamic variables which change with the placeholder exams from year to year. Some of the dynamic variables discussed in this paper include Age, Effective Class and Months Contributed. The age of the airman is recalculated and stored for easy retrieval at various points throughout the year. Effective Class records the medical class certificate held by the airman on the last day of the year, while Months Contributed accounts for the total number of months the airman contributed as an active airman for the given year. Months Contributed can range from 1 to 12 months for any specified year.

RESULTS

The data reengineering allowed the creation of a very large denormalized dataset in the format required by the statistical and data mining programs used by the Bioinformatics Research Team. Examination of the counts of active airmen by year revealed an anomaly in the numbers of electronic medical certificates issued during the years 1994 through 1999. Roughly 50% of the electronic medical exam records in this time period omitted the medical class issued for the certificate. This caused a large dip in the count of active airmen for this time period. Legacy data archived before the implementation of DIWS was used to extract and restore these lost records correcting the dip during this period. Figure 1 shows the counts of active airmen, by year before and after the correction of the data anomaly, and highlights the dip in pilot counts from 1994–1999.

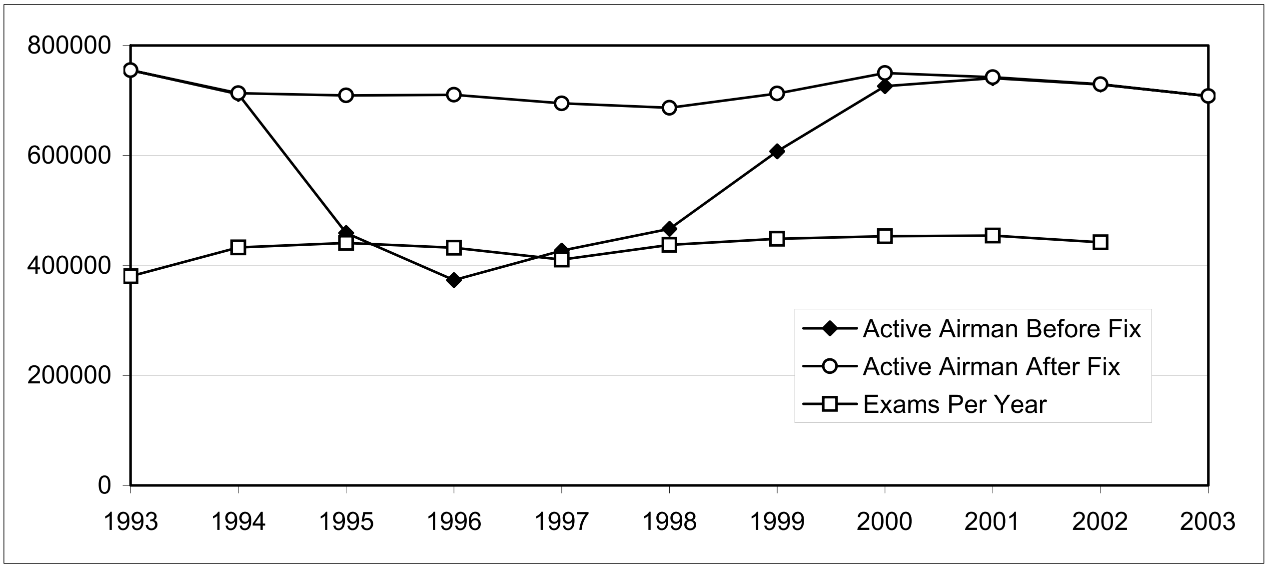
This restoration resulted in the inclusion of an additional 1.4 million exam records, from slightly more than 425,000 distinct airmen. This inclusion of medical records, corrected solely by the determination of their correct historical medical class, had the effect of discovering additional accident records. Further publications will elaborate on the epidemiological findings and modeling efforts of this data.

DISCUSSION

In order to achieve a foundation for aviation-related epidemiological studies using the entire U.S. pilot population, we have created a Scientific Information Management System (SIS).

This foundation will allow for the study of the distributions of the various pathologies within the U.S. pilot population and their impact on aviation safety. Rule changes within aviation medicine can be studied within the context of the entire pilot population, allowing for a detailed analysis of the results of those decisions.

Figure 1 – Number of Active Airmen by Year (Pre and Post Correction of Data Anomaly)



A complete examination of the evolution of the U.S. pilot population can be made allowing for a discussion of how the FAA's customer base is changing over time. Models can be constructed to predict future changes within this population and the possible regulatory actions that might need to be introduced to handle future requirements within aviation medicine. Since the Scientific Information System is comprised of data from multiple databases, these models should reflect changes introduced from these different areas (i.e., NTSB, Airmen Registry, DIWS, etc.). Models such as this will allow the observation of the combined affects of decisions made in different functional areas on the pilot population as a group. The SIS can serve as launching point for a large number of epidemiological studies that contribute to aviation safety and add to our knowledge of the ever-changing U.S. pilot population as a whole.

The data files are archived on the SIS server after the publication of each study. Archiving the data used for research projects within the SIS allows for the re-examination of study data years after completion of the study, when the original data sources that are incorporated into the AAMD-DSS warehouse will have changed in size and also will have ongoing quality changes made to their datasets. For instance, the NTSB publishes update files for its aviation mishap dataset that can include modifications made to older accidents for which updated information has been determined.

This process will permit compliance with the FAA-adopted guidelines that resulted from the 2002 Data Quality Act, also known as the Information Quality Act. This act was enacted as Section 515 of the Treasury and General Government Appropriations Act of 2001 (PL 106-544, H.R. 5658). The section directs the Office of Management and Budget to issue government-wide guidelines that "provide policy and procedural guidance to Federal agencies for ensuring and maximizing the quality, objectivity, utility, and integrity of information (including statistical information) disseminated by Federal agencies."

The major limitations of this complex aeromedical research effort are found in the data quality issues related to both the pilot electronic medical record and the electronic mishap records. Improvements in and added detail to the aviation electronic records will allow more complete studies to determine the factors with the most impact on aviation safety. The descriptive and predictive value of our modeling efforts will also benefit from improvements in the amount and quality of information in the input electronic records.

Showing AMEs, regulatory officials, accident investigators, and the general public that previously unknown knowledge can be found in the millions of existing records may help show all concerned the value of diligently applying the medical certification standards and the value of thorough mishap investigations.

CONCLUSION

The pilot electronic medical record, long a feature in the FAA handling of medical and other certificate records, along with the electronic aviation mishap record, can finally be analyzed using modern epidemiological methods. This research supports the medical certification decision-making of the regulatory component of the FAA Office of Aerospace Medicine. This is a permanent Scientific Information System for performing epidemiological studies concerning aviation medicine.

In future reports, we will be examining other derived variables, such as a measure of pilot experience, created by factor analysis from components of the SIS, to compare with months-contributed and other important variables to explore the important relationships between them and the improvement of aviation safety for both the private and the professional pilot, as well as their crew and passengers.

REFERENCES

Dunham, M.H. (2003). *Data Mining : Introductory and Advanced Topics*. Upper Saddle River, NJ: Prentice Hall/Pearson Education.

Hoaglin, D.C., Mosteller, F., & Tukey, J.W. (2000). *Understanding Robust and Exploratory Data Analysis* (Wiley Classics Library ed.). New York: Wiley.

Holt, R.W. (2001). *Scientific Information Systems*. Aldershot, England; Burlington, VT: Ashgate.

Kay, E.J., Harris, R.M., Voros, R.S., Hillman, D.J., Hyland D.T., & Deimler, J.D. (1994). *Age 60 Rule Research, Part III: Consolidated Database Experiments Final Report.* Federal Aviation Administration Office of Aviation Medicine Technical Report No. DOT/FAA/AM-94/22.

Whinnery, J.E. & Slaughter, J.W. (1981). A Computerized Repository System for Collection and Retrieval of Acceleration Stress Research Data. *Methods of Information in Medicine, 20*(1), 10-15.

APPENDIX A

Definitions

We define active airman to mean the holder of a pilot certificate who also holds a valid medical certificate. Figures A.1 thru A.8 illustrate the sequential algorithms used to collect active airmen records (pilots with a valid medical certificate) and create groups of interest for specific epidemiological research projects in a given time frame. Tables A.1 and A.2 provide the definitions and acronyms used in the flow charts and specify the data sources that are being queried at various phases in the construction of the SIS. The flow chart key attached to the first figure explains the shapes and symbols used in the charts. Figures A.1 thru A.6 also explain the sequence of decisions that determine whether candidates are members of the U.S. civil pilot population. Figure A.7 is a process flow chart recounting how pathology codes were assigned. Any pathology code of interest to aviation medical researchers could be identified for study. Figure A.8 describes how the dynamic variable Effective Class is created, which gives the class of medical the airman holds at the end of the year.

Table A.1 – Flowchart Definitions and Acronyms

AAMD-DSS	The Aviation Accident and Medical Database-Decision Support System.
Active Airman	An airman who holds a current medical certificate for any part of the year under study.
AIDS	The Federal Aviation Administration's Accident and Incident Data System.
DIWS	Document Imaging Workflow System located at CAMI which contains the electronic records of all pilot medical exams.
Effective Class	Medical Class, by itself, is a dynamic variable which changes over time. Effective Class is the computed medical class held by each airman on the last day of the given year.
Expiration Date	The expiration date of the medical certificate as recorded in the electronic exam record.
Expire Date	The results of the algorithm which calculates, from the electronic record exam date, an expiry date based upon age of the pilot and standard rules for duration of medical certificates in effect at the time of the exam.
EY	This is a variable that represents the end year of the study.
Months Contributed	The number of months the Active Airman held a current medical during the current year.
NTSB	National Transportation Safety Board.
SIS	Scientific Information System for Aerospace Medical Research.
SY	This is a variable that represents the start year of the study.
YYYY	This is a variable that represents the current year of the study.

Table A.1 – Flowchart Table Descriptions

Table Name	Description
Table A	The table that contains all of the NTSB events for a given time frame.
Table B	The table that contains the AIDS data for the given time frame.
Table C	The table that contains the medical information for the potentially active airmen for the given time frame.
Table D	The medical certificate information for each of the potentially active airmen.
Table E	The product of the join between the airman certificate data (Table D) and the airman exam data (Table C).
Table F	This table contains the records from Table E that fall within the maximum possible time frame that define our active airmen for the given year, which is based upon the medical certification rules at the time.
Table G	The records, pulled from Table F, which contain the most recent exam for each individual airman.
Table H	The product of the joining of records between Tables F and G. There will be a separate Table H for each year used in the process.
Table I	If an airman had more than one exam on the same day, we removed those exams from our dataset. Table I is the same as Table H with these multiple exams removed.
Table J	The Active Airmen for each year were appended together into one table. Basically, Table I is merged from each year into a single construct.
Table K	The table that contains all the selected pathology data.
Table L	This table contains records from Table K where the selected pathology was coded as a current condition.
Table M	This table contains the records from table K that document an airman's first occurrence of the selected pathology.

Figure A.1 – Active Airman Algorithm

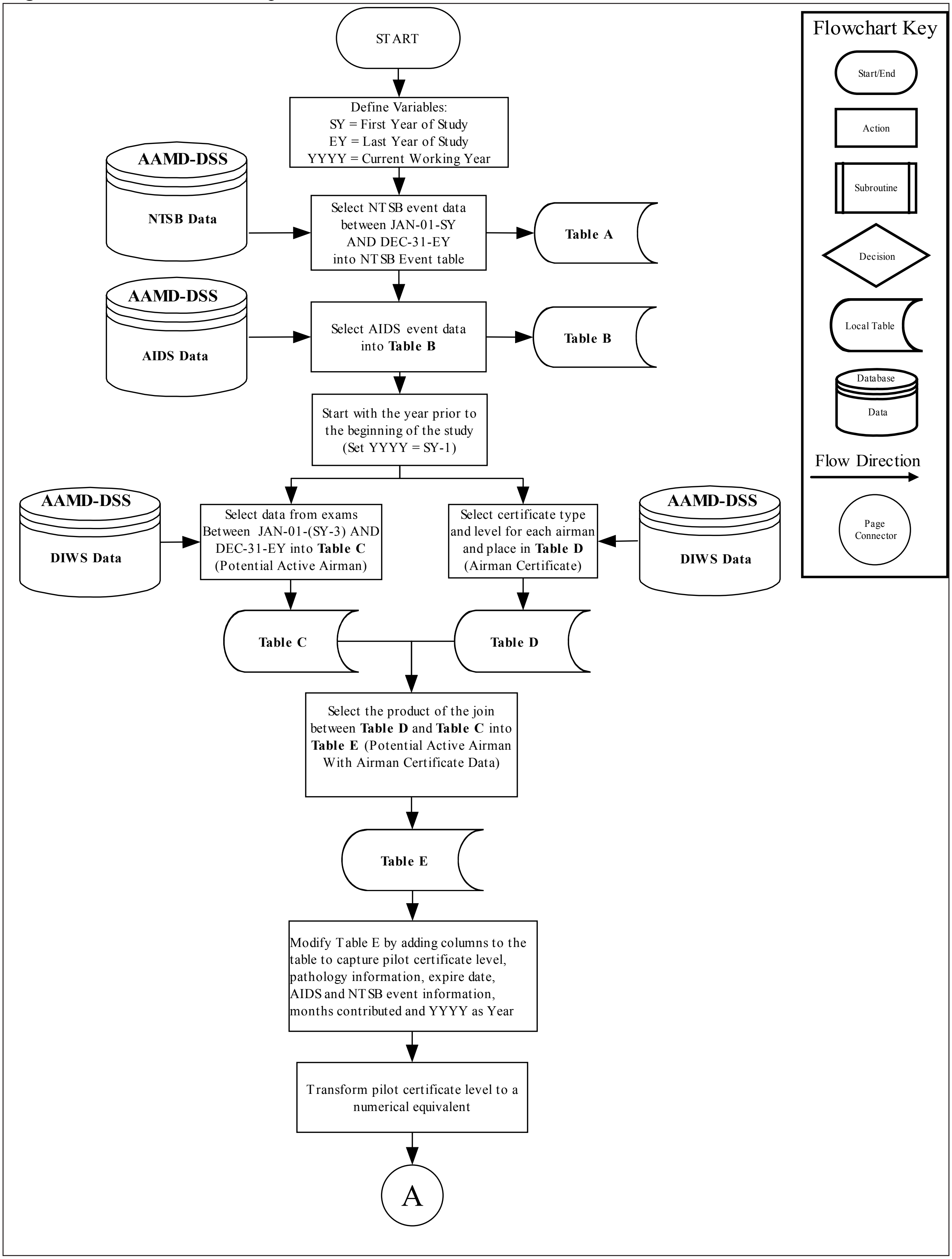


Figure A.1 – Active Airman Algorithm (Continued)

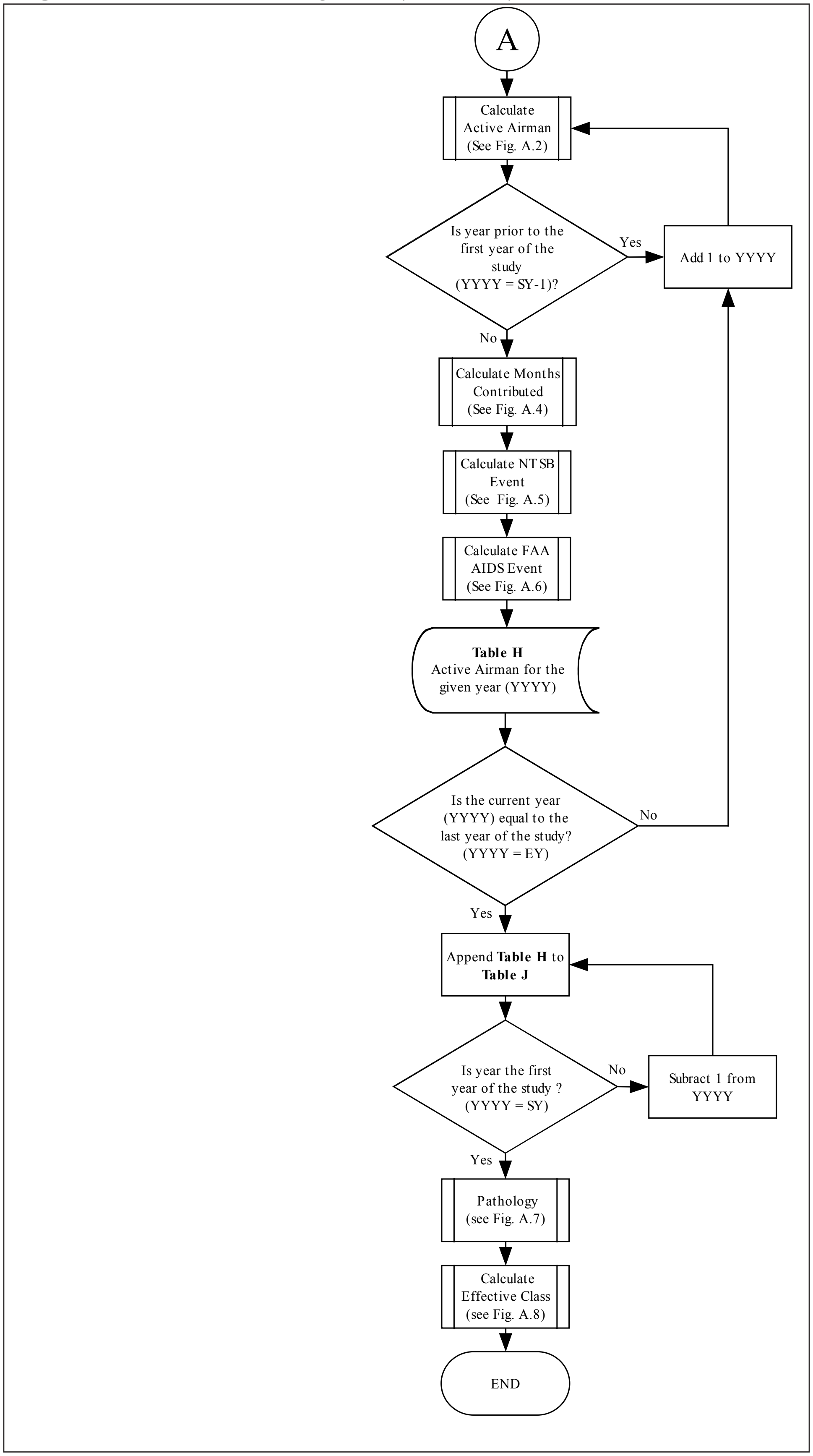


Figure A.2 – Active Airman Status Determined

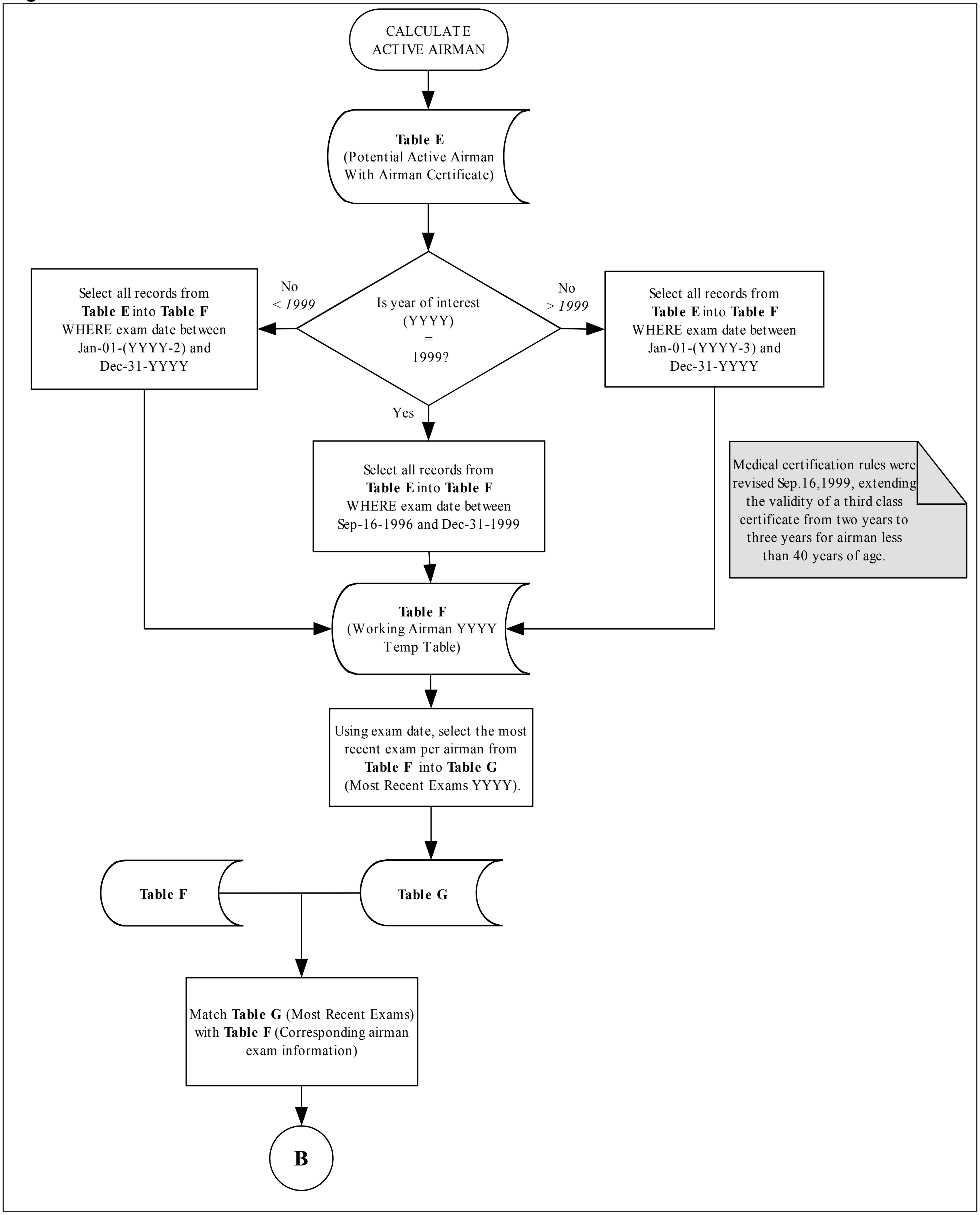


Figure A.2 – Active Airman Status Determined (Continued)

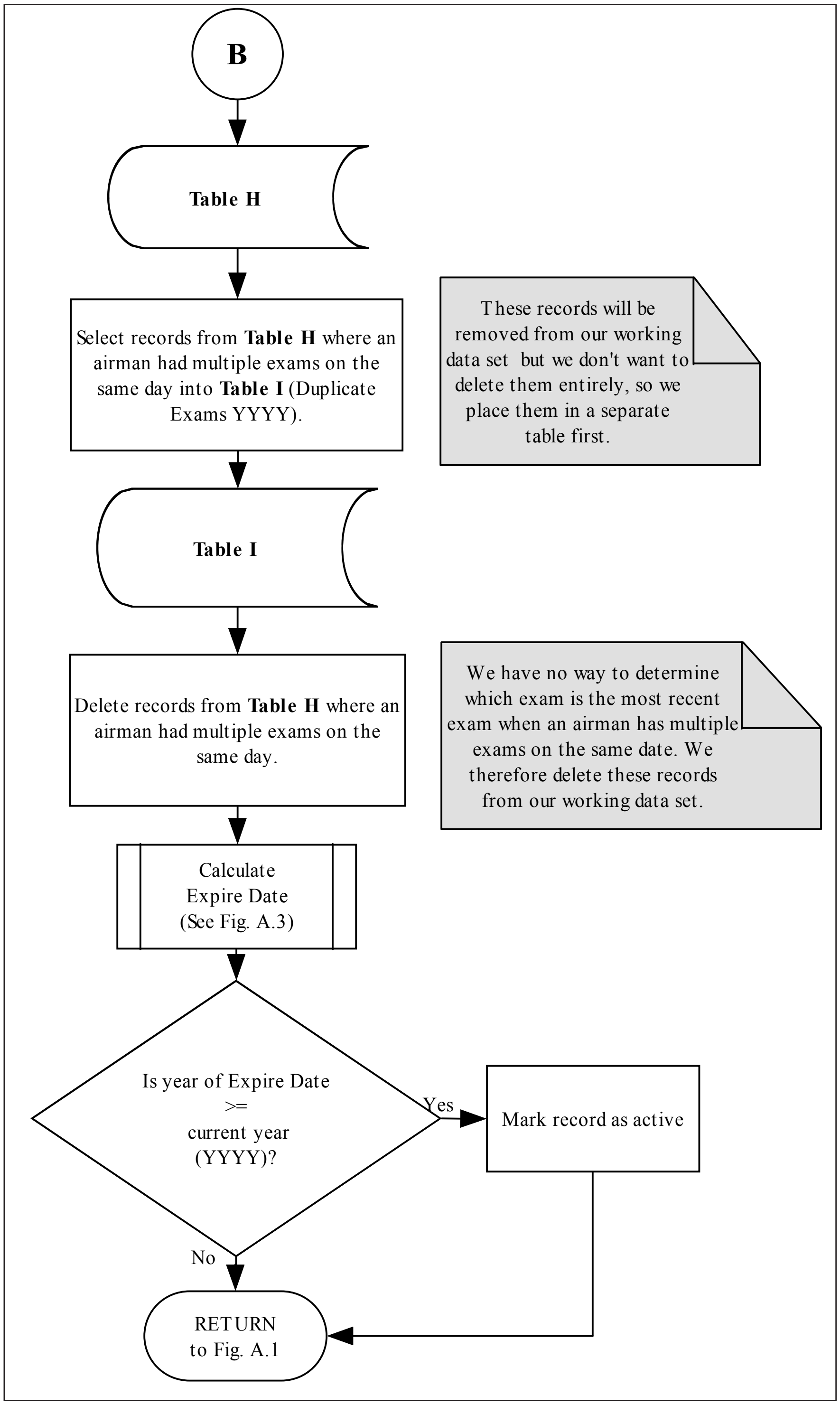


Figure A.3 – Calculate Expiry Date for Each Valid Medical Certificate, Account for Time Limited

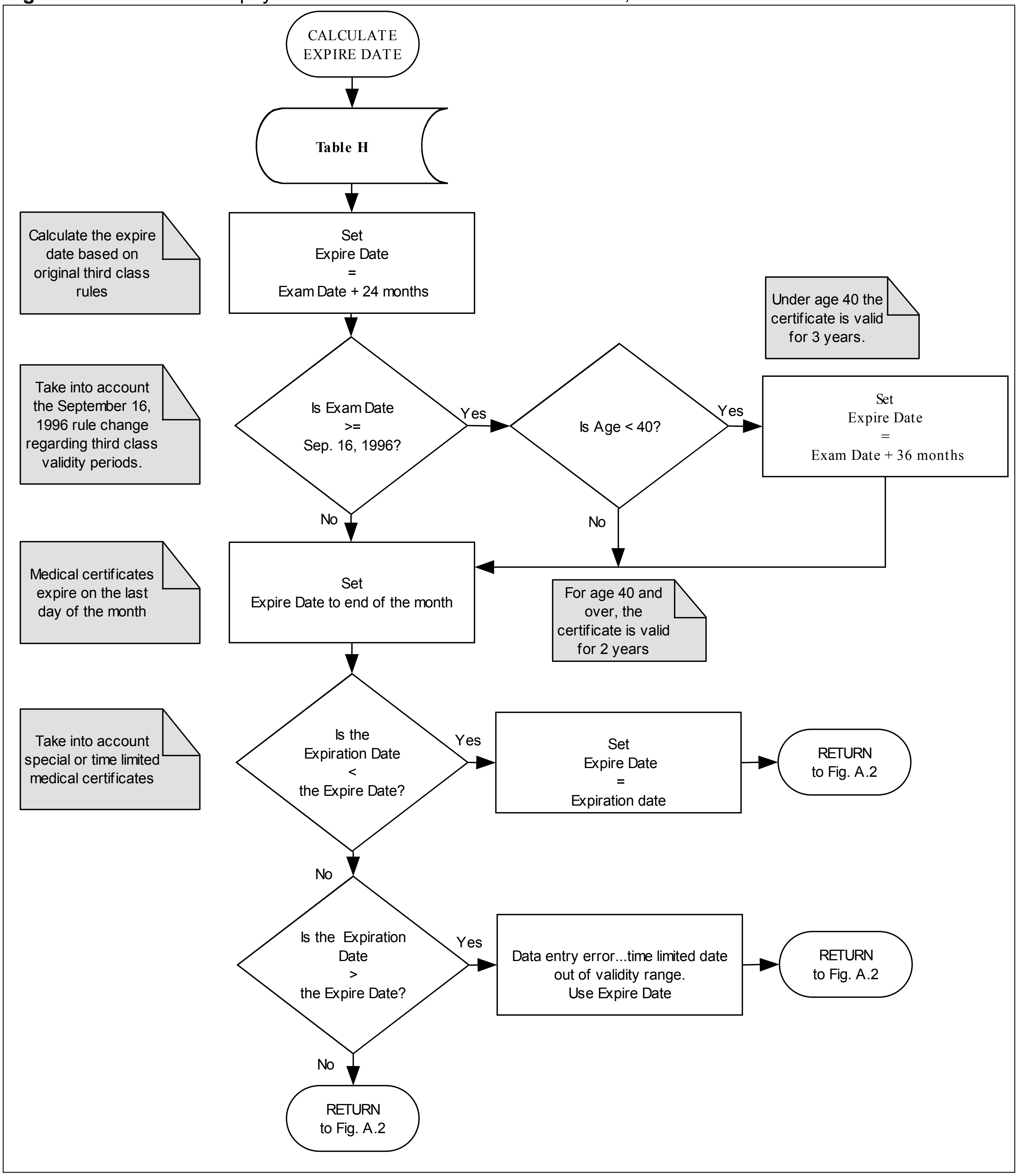


Figure A.4 – Calculate Months Contributed by Valid Certificate in Each Calendar Year

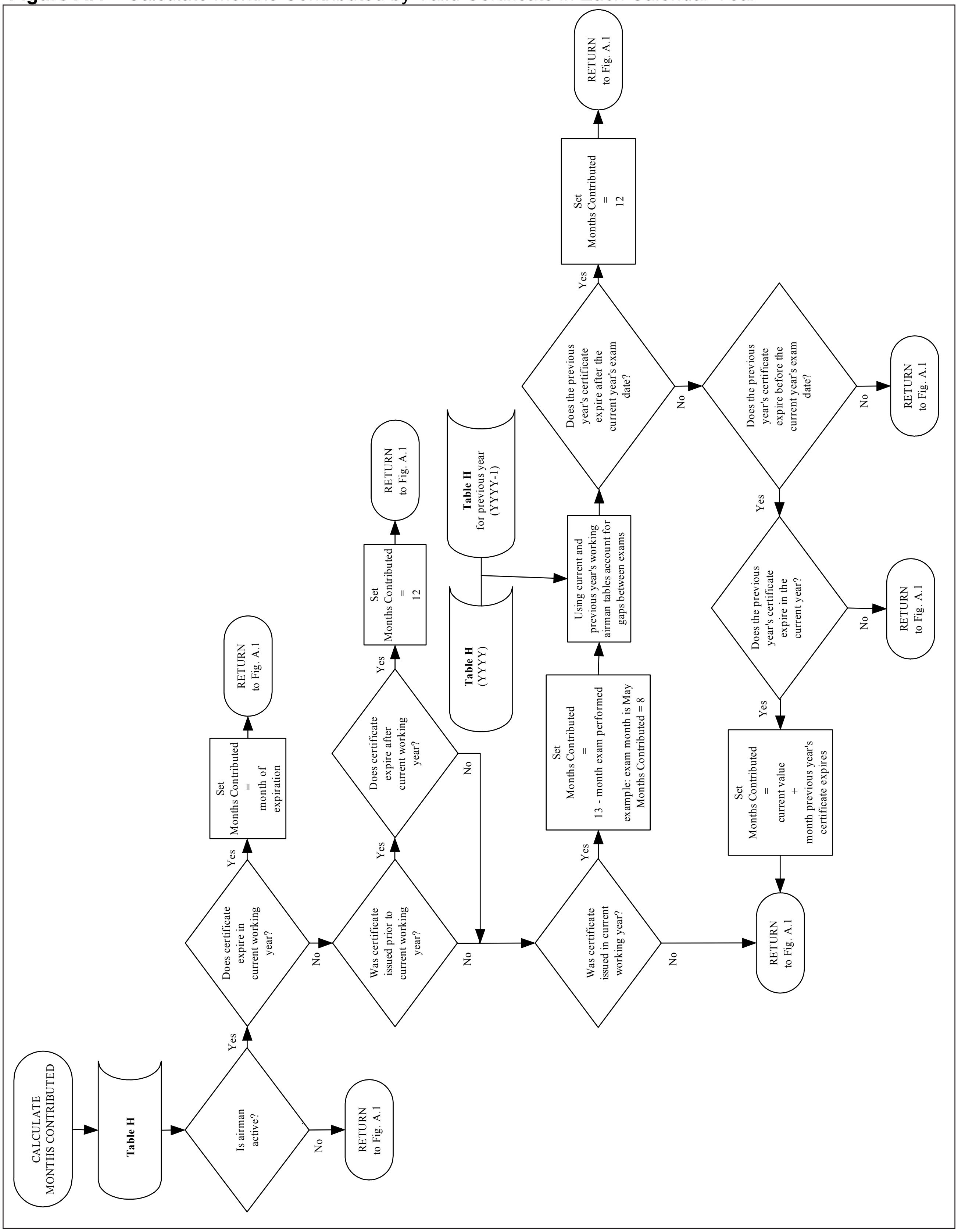


Figure A.5 – Determine if an Airman Was Involved in an NTSB Event During Current Calendar Year

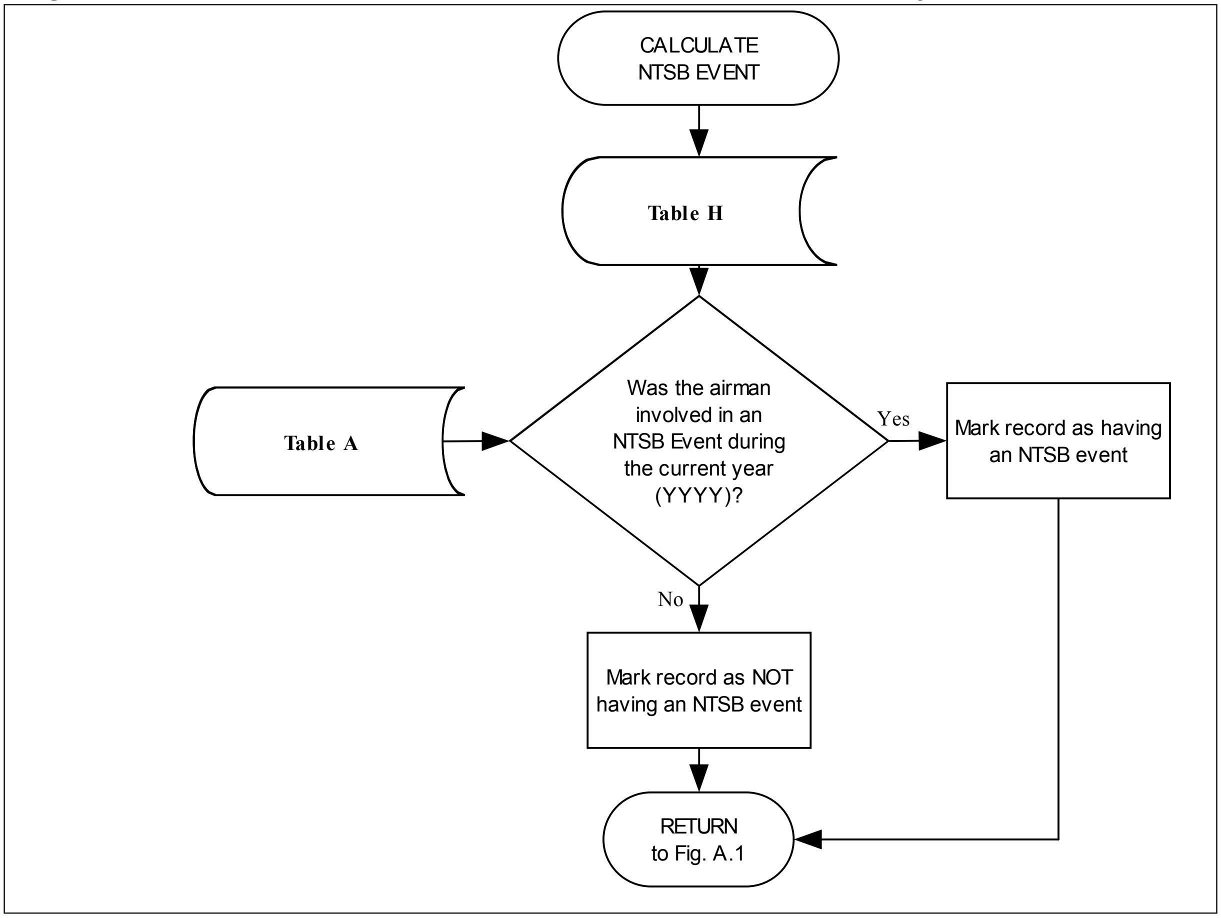


Figure A.6 – Determine if Airman Incurred FAA AIDS Event During Current Calendar Year

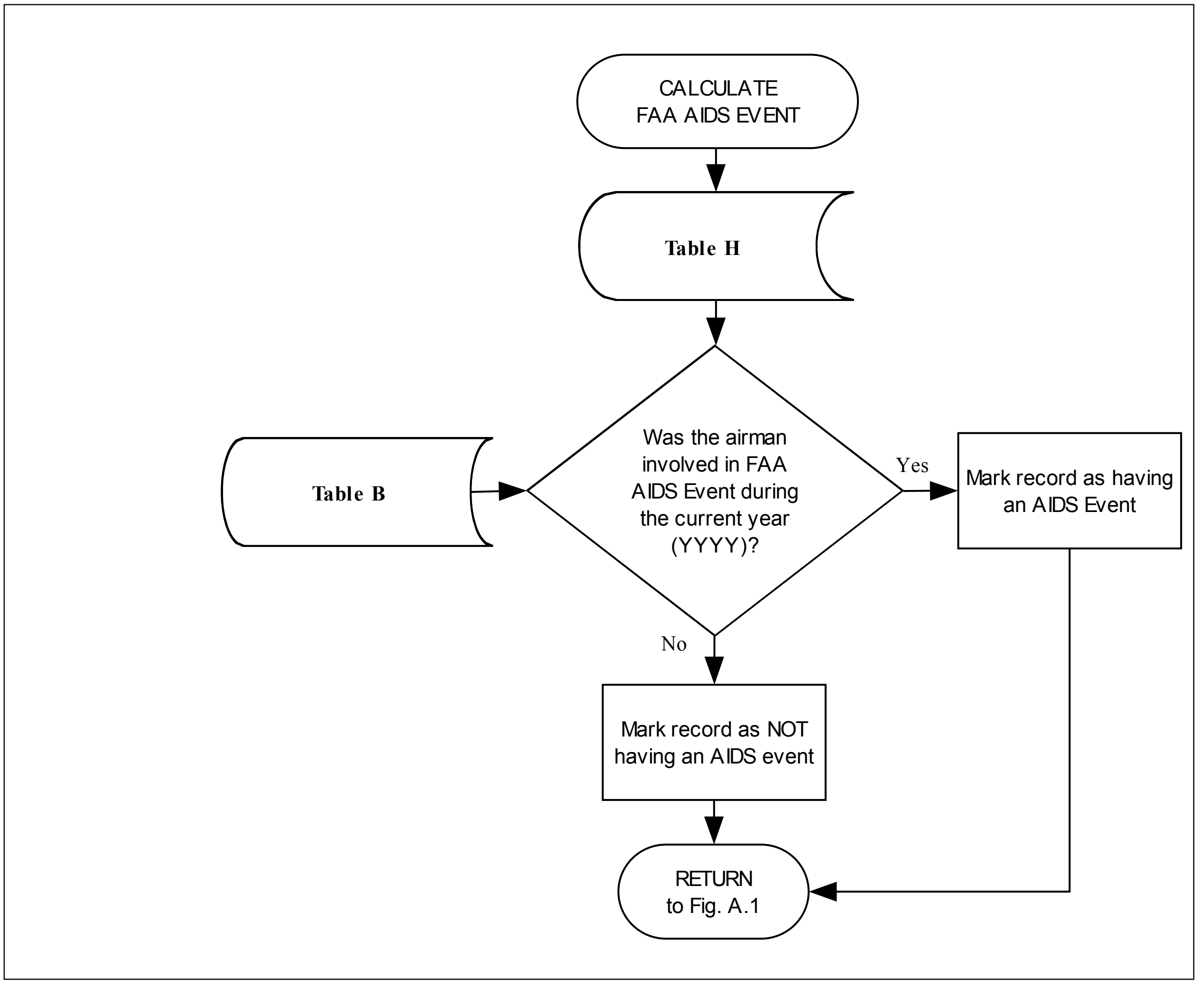


Figure A.7– Pathology Code Algorithm (Used To Select Pathology Data For Study Groups)

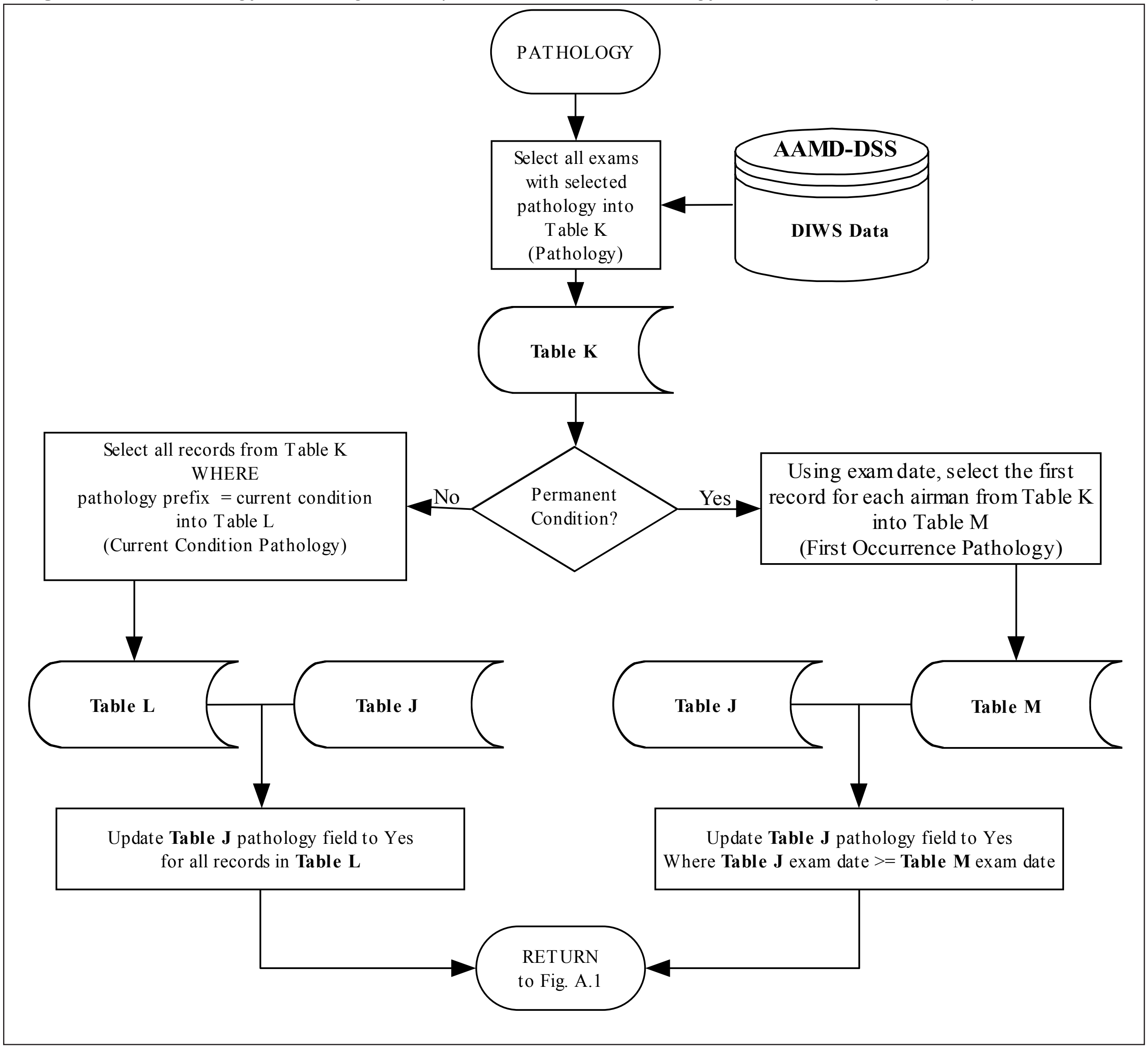


Figure A.8 – Effective Class Dynamic Variable Algorithm (to Determine Each Airman's Medical Class at Year End)