Journal of Athletic Training 2023;58(7/8):655–663
doi: 10.4085/1062-6050-0063.22

www.natajournals.org

Pac-12 Health Analytics Program: An Innovative Approach to Health Care Operations, Data Analytics, and Clinical Research in Intercollegiate Athletics

Kevin C. Robell, MA, ATC*; Marc F. Norcross, PhD, ATC†; Adam D. Bohr, PhD‡; Kimberly G. Harmon, MD§

*ARA Medical Enterprises Inc, San Diego, CA; †School of Biological and Population Health Sciences, College of Public Health and Human Sciences, Oregon State University, Corvallis; ‡Department of Integrative Physiology, University of Colorado Boulder; §Department of Family Medicine, University of Washington School of Medicine, Seattle

The objective of this study was to describe the purpose, methods, and effects of the Pac-12 Health Analytics Program (HAP) approach on sports medicine informatics, research, analytics, and health care operations. Sports injury-surveillance initiatives have been supporting the clinical research community in sports medicine for nearly 4 decades. Whereas the initial systems tracked only a few sports, current surveillance programs have expanded to include entire professional and elite athlete organizations, providing important statistics on sports injury risk management. The HAP is a conference-wide data-sharing and-analytics program. It collects authorized, deidentified clinical data, encompassing multiple domains of sports medicine injury management, including sports injuries and illnesses, concussions, risk exposure, and COVID-19 testing elements. The HAP provides clinicians with access to curated data to inform evidence-based practice and support local health care operations with respect to emerging sports injury trends. The HAP supplies approved research groups with access to a data repository that describes a homogeneous, elite intercollegiate athlete sample, thereby supporting nonresearch clinical initiatives as well as contributions to peer-reviewed research that can improve the health and well-being of Pac-12 student-athletes. The HAP is a novel approach to sports injury epidemiology and surveillance that has allowed the Pac-12 Conference to meet larger objectives regarding improving the student-athlete experience and clinical research among its member schools. Data quality control has improved the accuracy of the data and value to clinical athletic trainers within the conference. Curated dashboards displaying aggregated project data offer clinicians data-driven decision-making tools that help inform sports injury risk management. As of 2021, the HAP had supported more than 3 dozen data requests. These investigations have resulted in numerous peer-reviewed research contributions to the sports medicine community with findings that have great potential to improve the health and well-being of Pac-12 student-athletes.

Key Words: data analytics, clinical research, sports injury surveillance

Key Points

- Advances in cloud software and technology provide resources for athletic training organizations to innovate sports injury surveillance and health care analytics.
- The growth in clinical data documentation and collection offers a tremendous opportunity to expand the value proposition for clinical athletic trainers as key health care providers.

In 2013, the Pac-12 Conference and its member universities launched the Pac-12 Student-Athlete Health and Well-Being Initiative (SAHWBI) to "find ways to reduce injuries, share current best practices and latest studies, and conduct research to uncover new ways to keep student-athletes as safe as possible."[1] A part of this initiative was a specific desire to support evidence-based clinical practice and policy development via the development of a comprehensive Pac-12 Health Analytics Program (HAP) that would (1) guide health care operations by monitoring for emerging injury and illness trends in the conference, (2) provide member institutions with regular comparisons of local injury and illness trends relative to the rest of the conference, (3) objectively examine the effect of rule changes or programs on injury and illness rates, and (4) supply a database for researchers to answer questions important for the health and well-being of all student-athletes. Given that achieving these objectives would require a data-collection and-management infrastructure that went beyond that in existing injury-surveillance systems (ISSs), the conference sought to develop and implement a comprehensive health analytics program across all member institutions.

Sports ISSs have evolved since the introduction[2] of the National Athletic Injury/Illness Reporting System in 1975.

As of 2016, a total of 15 distinct ISSs were cited in the literature.[5] In general, the objective of sports ISSs is to assess longitudinal trends in injury incidence and characteristics among large populations such as collegiate athletes. They are intended to understand injury trends broadly; this information is critical to informing the need for and effects of rule, equipment, and coaching-technique changes.[3,4,5] In many cases, these systems are more of a "big-picture" approach to injury surveillance.

Whereas the National Collegiate Athletic Association (NCAA) Injury Surveillance Program (ISP)[5] currently operates as the national standard for sports injury surveillance among intercollegiate athletics and is perfectly designed to meet these aims, the ISP—as well as other existing surveillance systems at the time—exhibited several differences that precluded the Pac-12 Conference from using them to achieve the broader clinical objectives of the SAHWBI. First, most ISSs were not designed to provide the sports medicine clinician with local, peer-comparative data that could be used at the point of care for internal health care operations and exploratory data analytics. These data sets are typically supplied in very large and otherwise raw formats, often requiring intense data cleaning and analysis before insight can be gained. Second, the intent of the NCAA ISP is to provide insight into sports injury from a national, representative sample across all 3 competitive NCAA levels. As such, injury trends captured in the ISP may not reflect the Pac-12's relatively homogeneous, elite, intercollegiate athletic population from the western United States. Finally, the sports medicine and athletic training industry has seen a dramatic increase in both the volume and types of data used in the prevention and rehabilitation of sports injury and monitoring of student-athlete wellness.[6] However, many ISSs, including the ISP, were developed to focus exclusively on the epidemiology of sports-related injury and illness and may not include some information on clinical management of those conditions.[3,7,8] For instance, the Big Ten-Ivy League Concussion Initiative prospectively followed student-athletes across many sports for important outcomes of sport-related concussion (SRC).[8] Yet this investigation was limited to variables addressing SRC and exposure data. To meet the Pac-12 SAHWBI objectives, the conference leaders recognized a vital need to create a comprehensive health analytics program in which injury surveillance is part of a larger infrastructure that encompasses and links to other essential areas of clinical sports injury management, such as SRC, communicable disease, sleep data, and mental health. Moreover, this program should support both clinical operations and facilitate research to enhance patient outcomes and health care operations by delivering vital information on emerging organizational trends in sports medicine.

The purpose of our report was to present the approach to health analytics taken by the Pac-12 Conference. Specifically, we describe the features and functions of the HAP that have allowed the conference to pursue the clinical objectives of the SAHWBI and could be used by other organizations to expand or create their own health analytics programs.

METHODS

Organizational Structure

The Pac-12 SAHWBI is funded by a collective effort of the Pac-12 and member universities to improve the health and wellness of student-athletes. The SAHWBI serves as the primary platform for all research activities involving the health, safety, and overall well-being of Pac-12 student-athletes. In the first year of this initiative, the Pac-12 established the HAP with the intent that it would serve as the primary sports injury registry and database for all member institutions. The HAP was established around a centralized documentation infrastructure with all participants using the same Health Insurance Portability and Accountability Act (HIPAA)–compliant, cloud-based electronic health record (EHR). Nine Pac-12 institutions transitioned from their existing EHRs to the HAP in 2016, with the final 3 doing so by the fall of 2020.

The HAP functions under the guidance of the Pac-12 SAHWBI board of directors. A project manager at the conference level and 2 campus administrator (CA) co-leads, who also serve individually as their school's local CA, provide administrative and clinical oversight to participating institutions to optimize data collection, data quality, and reporting. Staff athletic trainers (ATs), as well as other health care providers where applicable, are required to supply initial and serial documentation records to the EHR platform. This requirement is essential to maintain quality data-collection and quality control initiatives. The CA co-leads, currently serving as senior staff ATs, assist in administering standardized clinical workflows and data-quality standards across the HAP. As an example of these standard workflows, every staff AT uses a standardized injury and illness evaluation form. The CA ensures that clinical staff know how to complete these forms, including documenting all HAP variables. In addition to HAP leadership, HAP documentation by staff ATs is overseen by an assigned CA located in the athletic department at participating institutions. At present, 10 of 12 CAs are staff certified ATs; the other 2 CAs hold various health care provider credentials or have a data science background. The CAs offer local leadership and oversight of clinical athletic training staffs to ensure accurate data entry and are also responsible for implementing weekly data-quality checks (please see the "Data Quality" section for more information). Other CA responsibilities include managing compliance with clinical documentation, daily quality control monitoring of HAP data, and interacting with other campus researchers and health care providers regarding future HAP data requests. The EHR platform has been standardized across the participating member schools to include identical case-report forms and workflows for clinical injury documentation; for example, completing the initial injury evaluation. Participating member schools are only allowed to access their own clinical data as they pertain to injury and illness management.

All HAP data are deidentified per the HIPAA Safe Harbor Law (45 CFR §164.514). The EHR vendor applies relevant deidentification strategies to all medical data residing in the system, including applicable HIPAA and Federal Educational Rights and Privacy (FERPA) considerations as well as student-athlete authorization, before the data are moved to a deidentified repository, the Pac-12 Research Portal. This research portal is a separate cloud database of HAP project data that is managed by the EMR vendor. The research repository is then made available to specified HAP leadership to achieve the objectives of the SAHWBI. (A more detailed explanation appears in

subsequent paragraphs.) Institutions have access to their own clinical data and can compare their local injury and illness information with conference norms via access to aggregate dashboards. Finally, information in the research portal is available to Pac-12–associated researchers via a standardized research request process after review and approval of the researcher's protocol by their institutional review board and the execution of individual data-use agreements.

Definitions

The HAP components detailed in this article provide a large data repository addressing many dimensions of sports injury and illness. Important operational definitions that facilitate standardization across participating member schools are as follows:

Injury and Illness. Sports medicine clinicians at participating member schools are required to report clinical information regarding HAP data elements. An *injury* or *illness* is defined as any physical or psychological condition affecting a student-athlete who is evaluated by (a) member (s) of the sports medicine staff and resulting in a clinical diagnosis requiring treatment or further evaluation. This definition, which is similar to that adopted for non–time-loss injury by the NCAA ISS in 2009–2010, was chosen because it allows HAP to include all injuries and illnesses, regardless of time loss, when the workload burden on staff ATs is analyzed.[5]

Relation to Sport. Both sport and non–sport-related injuries are captured in the HAP. An injury or illness that occurs "directly as a result of participation in a NCAA or Sport/Team–sanctioned event" is classified as a *sport-related injury or illness*, whereas non–sport-related injuries and illnesses have been documented as occurring outside of NCAA sanctioned activities.

Orchard Sports Injury Coding System. All clinical diagnoses documented in the HAP use a modification[9] of the Orchard Sports Injury Coding System (OSICS), version 10.0. Among popular diagnosis coding systems such as SNOMED or the International Classification of Diseases–10, the OSICS, version 10.0, includes diagnosis nomenclature that is sports medicine-centric when compared with existing technology.[10] Furthermore, this modification allows HAP leadership to add diagnoses frequently used in the clinical setting that represent evolving clinical practices at participating member schools. Updates to the OSICS are provided on a biannual basis and are based on clinical feedback and subsequent needs analysis.

Health Care Operations. Adapted from 45 CFR §164.502, *health care operations* in the HAP are defined as non–research-based, organizational quality-improvement activities designed to enhance patient outcomes and mitigate the risk of sports injury. An example of these activities may include end-of-year injury reporting, the incidence of lower extremity injury incidence across event types, referral patterns for patients with foot and ankle conditions, or weekly injury incidence trends across a competitive season. These activities are supported by a technology infrastructure, in both software and cloud resources, designed to help identify emerging sports injury trends occurring within the sports medicine organization.

Time Loss. Clinical documentation of missed or lost time is managed for each injury or illness. Days lost from injury are calculated by using the difference between the date a patient with an injury or illness was classified as "out" and the date he or she was cleared for unrestricted physical activity.

Advanced Care Workup. Along with collecting various sport-related injury dimensions, variables that generally describe any advanced care provisions are also integrated in the project. The following Boolean variables (true or false) are included: Did the injury or illness require a follow-up visit with a team physician? Did the injury or illness require a diagnostic imaging study (eg, radiographs, magnetic resonance imaging)? Did the injury or illness require surgery? Did the injury or illness require the use of a prescription medication? Did the injury or illness require an advanced treatment procedure (eg, sutures, intravenous fluids, infection management)?

Pac-12 Research Portal. A secure, cloud-based research portal and repository stores authorized, deidentified components of the HAP in a structured format to be used for health care operations within the conference and to support research through secondary-use requests. The research portal is actively managed by the EHR vendor, and all deidentification occurs before the data are released to the HAP.

PacTrac Event Definitions. Exposure documentation workflows require the clinician to relate injury exposure to specific events. Operational definitions for the various types of events that are captured in the PacTrac exposure module are given in Table 1.

Domains

The primary domain of clinical data collection initially included approximately 30 variables addressing "sports injury and illness." In 2018, the HAP expanded to incorporate variables representing initial, serial, and return-to-play assessments included in the Pac-12 Concussion Assessment, Research and Education (CARE)–Affiliated Program (CAP) on SRC.[11] In the early years, injury metrics could only be expressed as *prevalence* (ie, the percentage of athletes with a condition) or incidence per athlete-season. In 2020, the HAP expanded to include athlete-exposure variables compiled in PacTrac, a daily event-exposure documentation platform designed to incorporate variations in exposure types and denominators, as well as logistical practicalities in collecting detailed exposure information.[12] In response to separate conference initiatives, the HAP further expanded in 2020 with surveillance testing[13] for COVID-19. These domains (Table 2) provide a range of clinical metrics to be reported as part of a health care operations paradigm, as well as used secondarily for research under approved data-sharing and data-use agreements. Although a detailed description of each data point in the HAP is beyond the scope of this communication, in total, more than 1000 variables related to sports injury and illness are captured in the HAP project.

PAC-12 Care–Affiliated Program

In 2014, the NCAA in collaboration with the Department of Defense (DoD) established the NCAA-DoD CARE

Table 1. Standardized Operational Definitions for Events Used in the PacTrac Exposure Module

Event Type	Operational Definition
Practice (team training)	A "traditional" practice activity in which an entire team of athletes engages in sport-related practice activities as a means to improve team performance.
Practice (walk-through)	A practice activity in which the only involvement of athletes is to rehearse or review specific plays or game situations. This activity may occur on the same day as a competition but does not include precompetition warm-ups.
Practice (individual skills session)	A practice activity in which athletes engage in sport-related activity as part of a small group with a member of the coaching staff at a time outside of a full-team training session.
Practice (scrimmage)	A practice activity in which the primary purpose is for competition in a sport between athletes on the same intercollegiate athletics team, such as an intrasquad spring football game or fall softball game.
Precompetition	Any physical activity performed immediately before a competition that is intended to prepare the athlete to engage in competition (eg, precompetition warm-up).
Competition	Any game, match, exhibition, scrimmage, or joint practice session with another institution's team, regardless of its formality.
Strength training	Any form of resistance training with the primary goal of increasing strength that occurs outside of a practice activity, regardless of location.
Conditioning	Any form of cardiovascular-focused training that occurs outside of a practice activity, regardless of location. This will most commonly be running but may also include agility and other fitness activities that would not primarily be considered forms of strength training.
Voluntary activities	A voluntarily attended event where the student-athletes are not required to report back to a coach or designee any information related to the activity and in which participation is not required.
COVID team-related pause	A fixed time frame, when all team-related activities (eg, practice, competition, strength training, conditioning) are suspended, specifically due to preventive measures under COVID-19.

Consortium (NCAA-DoD CARE Consortium).[14] This large-scale effort seeks to define the natural history of concussion by collecting core common data elements across more than 30 universities and US service academies. In August 2017, recognizing the substantial overlap in common goals of both the NCAA-DoD CARE Consortium and the Pac-12 SAHWBI, the Pac-12 collectively became a CAP—effectively serving as a regional hub for the CARE Consortium.[11] By sharing many of the elements of the parent CARE Consortium, the inclusion of the Pac-12 CAP data set significantly expanded the HAP to include more than 900 variables on the evaluation and management of concussion. The initial CAP data-collection methods were managed by the Pac-12 Concussion Coordinating Unit at the University of Colorado Boulder with oversight from the Pac-12 Brain Trauma Task Force. Components of the CAP were baseline concussion data, including medical and concussion history, the Standard Concussion Assessment Test, the Standard Assessment of Concussion, and the Balance Error Scoring System. In addition, baseline cognitive inventories (ie, ImPact, King-Devick), as well as vestibular and oculomotor measures through virtual reality headsets, were incorporated.

Table 2. Health Analytics Program Domains[a]

Component	No. of Variables
Sports injury and illness	
Demographic information	7
Injury or illness	9
Event	8
Outcome	7
Concussion	
Demographic information	11
Baseline concussion testing	357
Postinjury serial testing	315
Concussion history	14
Medical history	62
Initial injury examination and return-to-play form	185
COVID-19	
Real-time polymerase chain reaction test	10
Antigen test	18
Exposure variables (PacTrac)	
Event participant	21
Event	18
Roster	13
Total	1065

[a] Domains of sports injury and illness included in Health Analytics Program documentation workflows. Table 1 also describes the specific subcomponents of each domain.

COVID-19 Surveillance Testing Program

Beginning in the fall of 2020, and in response to the effect of the COVID-19 pandemic on collegiate athletics, the HAP expanded to include standardized data on real-time polymerase chain reaction and antigen testing protocols. As a separate health care initiative, daily Sofia2 antigen tests (Quidel) were processed and the results imported to the EHR platform to support athlete management and safe return-to-play decision-making. Authorized real-time polymerase chain reaction and Sofia2 antigen test results stored in the research portal are being used to sustain ongoing surveillance and contact tracing initiatives in the conference and were used to support prospective research conducted in the spring of 2021.[13]

Mental Health Coordinating UNIT

In the fall of 2020, the Mental Health Coordinating Unit (MHCU) was created by the SAWHBI board as a grant-funded initiative to establish mental health screening methods consistent with NCAA best-practice recommendations to be implemented at each Pac-12 university.[15] Led by the University of Arizona, the MHCU began contributing data to the HAP in the fall of 2021. Inclusion of the MHCU will considerably increase the capabilities of the HAP in improving student-athlete health and well-being.

Exposure Tracking (PacTrac)

Injury rates have been variously reported to facilitate comparisons among groups or time periods or to monitor trends over time.[16,17] The NCAA ISS/ISP has long defined *athlete-exposure* as 1 student-athlete participating in 1 NCAA–sanctioned practice or competition in which he or she was exposed to the possibility of athletic injury, regardless of the time associated with that participation.[5] However, injury incidence rates and the associated injury risk ratios calculated when conducting comparisons can be significantly influenced by the actual participation hours of of the individual athletes.[16,17] In 2020, the HAP widened its clinical data-collection operations to include PacTrac based on Pac-12 SAHWBI-funded research and development.[12] PacTrac minimum reporting standards were adopted to include all "countable athletically related activities" (CARA) as defined by bylaw 17 (17.02.1) of the NCAA Division I Manual.[18(p 241)] Although clinical documentation and resulting analytics are feasible within PacTrac for non-CARA activities, formal standards for these activities may create an additional burden on clinical staff to report participation, and therefore, they were not included in the minimum standards. Furthermore, according to an individual feasibility assessment of each sport in the conference, sports were assigned to 1 of 3 tiers on the basis of the ability to capture granular exposure metrics at the individual athlete level and associated medical coverage models.[12] Granular metrics used for tier 1 sports were assigned on the basis of identification of the most detailed unit of elapsed time for each event (ie, plays, minutes, points, events) that could be feasibly captured by ATs at all conference institutions. Documenting exposures at these deeper levels offers a more accurate assessment of the individual and sport-specific injury risk.[16,17] To our knowledge, no precedent exists for an exposure collection platform acquiring this level of detail on daily athlete-exposures for individual student-athletes across multiple institutions across an intercollegiate athletic conference.

Tier 1 sports are generally sports that require a certified AT present at all CARA activities. Documentation requirements for these sports include daily tracking of the "individual time" a student-athlete participates in a CARA activity (Table 3). Granular metrics for competitions were used for tier 1 sports to further capture exposure as each individual athlete's relative contribution to a competitive event.

Tier 2 sports are generally sports that may not have a certified AT present at all times or do not have a consistent training plan for all athletes (eg, pitchers versus outfielders, long-distance runners versus sprinters). Accordingly, the collection of granular data in these sports is not feasible. Tier 2 sports include baseball, beach volleyball, cross-country, diving, rowing, softball, swimming, tennis, and track and field. The documentation standards for tier 2 sports are that ATs track daily athlete-exposures using the NCAA ISS definition—simply whether or not the athlete participated in the event—but they do so for each athlete individually.

Tier 3 sports are Pac-12–sponsored sports that traditionally do not (1) require direct clinical coverage or (2) have a practice or competition facility on campus. Traditionally, athletic training personnel are not informed of the daily activities in these sports and student-athletes may only access medical care on an as-needed basis. Tier 3 sports include acrobatics and tumbling, fencing, rugby, sailing, skiing, squash, and synchronized swimming. Daily exposures are not tracked for athletes in tier 3 sports. Rather, athlete-seasons are the exposure metric that could be used to calculate injury incidence rates.

Exposure documentation in PacTrac includes customized data-quality mechanisms to ensure high-fidelity data are captured both at the clinical level and in the research portal. Clinical documentation efficiencies were implemented to further ensure high-quality data: (1) multiselection of athletes, (2) individual exposure estimation, and (3) duplication of event and roster details. Bimonthly data-quality checks are submitted by CAs. Data-quality checks are aggregated and curated for future documentation strategies and maintaining low rates of missing data.

Data Handling and Processes

All clinical data collected by clinicians at participating member schools are subject to a series of steps intended to improve documentation and data quality. Data-quality workflows are managed by the HAP CA and conducted on a bimonthly basis. Submitted reports are reviewed for completion and accuracy.

Data Quality. Inherent challenges exist with any data project when establishing quality assurance or quality control mechanisms. Further validating these mechanisms presents a separate challenge. Managed locally by the CA, the HAP program has implemented unique and scalable data-quality mechanisms to ensure project data are accurate and complete. Two primary components of data-quality strategies, null data analysis and logic checks, were applied to all documented clinical data in the HAP. In 2019, the data-quality process was externally reviewed by IQVIA Inc, a global data-analytics firm specializing in analytics and epidemiology. Records that fail these comprehensive analyses are then corrected in the medical record by staff ATs.

Null Data Analysis. Nonexistent, missing, or null data can lead to bias and reduce the efficiency in a data set**.** Differing guidance exists regarding minimum and maximum thresholds of missing data.[19,20] The HAP program uses a 5% maximum threshold for missing data. This threshold was chosen to maximize the completeness of HAP data for secondary research use. Data-quality strategies were implemented to maintain this 5% null data rate. Null data rate analysis in the HAP program leverages programmatic reporting solutions circulated bimonthly to all staff ATs. Reporting strategies identify records with missing data in the health record that need attention. The HAP CAs manage this ongoing quality control process and supply a monthly aggregate null rate for further quality assessments. Customized system workflows and enhanced communication, as well as guidance to stakeholders, provide a rigorous and repeatable framework to correct clinical records containing missing data while improving the overall completeness and accuracy of clinical documentation habits. As of May 2021, the HAP had a 1.1% null data rate.

Logic Check Analysis. In addition to missing data, data points that are nonsensical when combined with 1 or more additional dimensions have been known to increase bias

Table 3. Competition Metrics

Tier	Athlete Contribution	Event Total
1		
Baseball[a]	Athlete contribution to innings	Total innings
Basketball	Athlete min	Total time, min
Field hockey	Athlete min	Total time, min
Football	Athlete no. of plays	Total No. of plays
Gymnastics	Athlete % of team events	No. of team events (eg, bars, floor)
Lacrosse	Athlete min	Total time, min
Soccer	Athlete min	Total time, min
Softball[a]	Athlete contribution to innings	Total innings
Volleyball	Athlete contribution to points	Total points in match
Water polo	Athlete min	Total time, min
2		
Beach volleyball	Participate yes (100%) / no (0%)	1 team scoring event
Cross-country	Participate yes (100%) / no (0%)	1 team scoring event
Diving	Participate yes (100%) / no (0%)	1 team scoring event
Rowing	Participate yes (100%) / no (0%)	1 team scoring event
Swimming	Participate yes (100%) / no (0%)	1 team scoring event
Tennis	Participate yes (100%) / no (0%)	1 team scoring event
Track and field	Participate yes (100%) / no (0%)	1 team scoring event
Wrestling[b]	Participate yes (100%) / no (0%)	1 team scoring event

[a] Denotes a hybrid sport that has adopted tier 1 standards for competition activities and tier 2 standards for noncompetition activities.
[b] Denotes a hybrid sport that has adopted tier 2 standards for competition activities and tier 1 standards for noncompetition activities.

and reduce validity in research data.[21] Consider the following example of 2 data points in initial injury evaluation: the incident date (ie, the date of injury) and the examination date (ie, the date the injury was evaluated by a staff AT). When assessing health outcomes of seeking care after concussion symptom onset, it would be nonsensical to have an examination date that occurred before the incident date. These types of data, if not addressed, may negatively affect results and create inconsistencies in the health record. Along with null data checks, systematic auditing of sensibility, or logic checks, have been implemented across the HAP quality control frameworks. A systematic approach to data quality helps ensure the appropriateness of clinical data across different injury scenarios and enhances the useability of project data for health care operations and secondary clinical research. As of May 2021, of the data points in the HAP, 97% passed established logic checks.

Programmatic logic checks are defined as data-quality checks performed as the clinician interfaces with the EHR and include (1) mandatory form controls and picklists, (2) clinical picklists and other controls that have been standardized across all participating schools and locked to prevent the addition of values, and 3) a programmatic process for the medical record that identifies failed clinical records and notifies the clinician. Programmatic logic checks are performed by the software application itself, similar to a level of intelligence that (1) informs the user in real time of data-entry errors or (2) standardizes the collection of HAP data that protects from data-entry errors.

Nonprogrammatic logic checks include data-quality standards that have been implemented to review clinical data that may have been overlooked in the initial programmatic phase. This phase of data-quality inspection involves weekly review and aggregation of clinical data by the CA to determine which records need correction. Information is passed back to the clinician to be verified and corrected in the medical record on the basis of the specific criteria listed earlier.[22]

PAC-12 Research Portal

Data Flow. Conference-level data-sharing agreements assign local clinical data ownership to participating institutions that grant authorized data sharing with the HAP. Before project data are made available for sharing and analysis, clinical information is stored and managed by the EHR vendor. As the primary ambassador for protected health information and data security, the EHR vendor stores and secures all clinical information per the end-user license and business associate agreements with each participating institution. In addition, data are processed and stored in compliance with all applicable federal regulations, including HIPAA and FERPA. At no time do the SAHWBI or the HAP project managers have access to identifiable clinical information stored in the secure EHR.

Data Deidentification. All project data stored in the research archive are subject to deidentification strategies. The main purpose of data deidentification is to reduce the risk of individual identity disclosure upon secondary use of HAP data for health care operations and research. Currently, all HAP project data are initially deidentified per the HIPAA Safe Harbor Law (45 CFR §164.514). Additional deidentification strategies are then applied to further reduce the risk of reidentification, including nearest-neighbor ("k-anonymity") rules, encrypted 1-way hashing of specific variables, and the inclusion of generalized hierarchies. These additional strategies further reduce the risk of reidentification on the basis of unique or rare injuries, unique sports offered only at individual schools, and data that describe a specific sport or school.

Furthermore, student-athletes at participating institutions must provide written authorization that allows sharing of deidentified data with the HAP. Authorization must be completed before their clinical data can be moved to the research portal. The HAP authorization is tracked in the campus medical record and managed by the CA. Only authorized data are allowed to move to the deidentified archive for analysis and sharing.

Health Care Operations and Exploratory Data Analysis. A distinct, separate infrastructure was created to supply the health care analytics and research components of the HAP. Governed and managed by the EHR vendor, HAP project data are refreshed quarterly. Deidentified clinical data are aggregated and distributed in the form of clinical dashboards using business intelligence and analytics software. Aggregated clinical dashboards are made accessible to participating member schools to be used for internal health care operations and end-of-year reporting, as well as within-conference comparisons across different injury and illness outcomes.

Secondary Use of HAP Data for Research. In addition to the health care operations and analytics utility, the HAP provides "approved research groups" in the conference with access to deidentified project data for secondary research use. Approved research groups are led by principal investigators who are directly or indirectly associated with the medical care of student-athletes in the conference. A comprehensive application process, data-needs analysis, and written documentation of local institutional review board determination are required.

Each HAP data-request application is reviewed by multiple members of HAP leadership and the SAHWBI board. Upon approval, individual data-use agreements are executed for each HAP request. Once a HAP request is approved, the resulting project data are encrypted and securely transferred to the HAP CA and then to the requesting principal investigator, subject to campus-specific procedures on research data. The HAP project data made available to approved research groups are refreshed quarterly. This unique characteristic of HAP data provides clinical researchers in the Pac-12 Conference access to updated and relevant data.

DISCUSSION

The Pac-12 HAP supports evidence-based clinical practice and policy and allows the conference to pursue the SAHWBI's clinical objectives by providing actionable feedback specific to Pac-12 athletes that conference clinicians can use to make data-driven decisions at different points of care across many clinical domains. Currently, the HAP contains more than 1000 dimensions of sports injury and illness across its 5 domains (sports injury/illness, COVID-19, PacTrac exposures, concussion, and mental health) and has followed over 15 000 student-athletes since 2016, equivalent to an approximately 90% authorization rate of all Pac-12 student-athletes. This sample gives researchers access to a relatively homogeneous sample of elite intercollegiate athletes that allows stakeholders to focus investigations on the specific needs and demands of these intercollegiate athletes in an effort to improve student-athlete health and wellness.

Whereas the sports injury/illness domain serves as the backbone of the HAP, the additions of CAP, PacTrac, COVID, and mental health variables (Table 2) are features that separate the HAP from traditional ISSs. Moreover, when coupled with rigorous data-quality processes and standards (<5% error rate), the HAP offers clinicians and researchers an expansive array of data that can be used to study a wide range of non–research-driven, clinical questions surrounding health care operations, clinical practice, and research. Examples of using HAP data for these types of non–research-based clinical initiatives include (1) evaluating weekly injury trends to better understand the clinician burden at different points during a competitive season in efforts to inform playoff expansion in collegiate football, (2) comparative analysis of hamstrings injury prevalence across a full athletic calendar, and (3) understanding emerging trends in AT referral patterns to physicians for patients with common lower extremity soft tissue injuries. More recently, the aggregated PacTrac has been used to monitor postinjury return-to-play data by providing the clinician with a breakdown of the actual versus expected workload. During the COVID-19 pandemic, clinicians were able to evaluate and track weekly positive testing rates of both surveillance or screening and contact-tracing protocols. Finally, HAP data were also used this past year to supply conference clinicians with an increased understanding of health care utilization patterns and possible inequities in student-athlete access to and characteristics of care. Whereas some of these queries were used internally to optimize patient care, operations, and the student-athlete experience, others were identified as pilot data for subsequent HAP research data requests for investigations that were ultimately submitted for publication.

The inclusion of CAP benefits other existing concussion surveillance initiatives by including an expansive set of standard of care assessments for SRC in a data ecosystem across multiple domains of care. This allows unique quality control and clinical research queries across multiple data sets to facilitate complex research questions on return to play and associated risk factors. The integration of PacTrac, the exposure tracking module, in standard clinical documentation workflows significantly enhances the ability of clinicians to collect the information so that it can be used for more robust health care analytics, such as measuring the injury incidence in sports sponsored by the Pac-12 Conference. At the height of the global pandemic, COVID features allowed clinicians to view real-time testing data, including weekly positivity rates. This allowed the SAHWBI board to inform conference athletic directors and other institutional leadership of relevant team-related pauses in activity and to monitor the resolution of isolations and quarantines. Biannual data collection of mental health screening questionnaire results allows mental health practitioners to intervene in a timely manner with individuals who demonstrate at-risk responses, as well as examine clinical reporting on epidemiologic outcomes. Collectively, these features of the HAP provide the infrastructure necessary to support both clinical and research practice.

The "big-data" age has fostered the creation of next-generation technology in cloud computing resources and efficient statistical and analytic software that, historically, was not only cost prohibitive for small and medium organizations but required significant commitments in software development and database management to implement and maintain. Existing cloud computing resources, business intelligence, and analytic software, along with machine learning and artificial intelligence frameworks, are becoming cost-effective approaches to organizational risk-assessment strategies. This big-data paradigm shift in health care operations has offered the sports medicine industry a

significant opportunity to benefit from warehousing, curating, and extracting insight from organizational data.[23] Organizations now have access to advanced data-management and-analytics technology with lower financial barriers to entry. This opportunity allows stakeholders to organize high-volume clinical data that may have been previously inaccessible or historically not included in return-to-play management of patients with athletic injuries. With the rapidly increasing library of therapeutic interventions that accompany advanced technology, future sports injury-surveillance initiatives need to account for these new data points and how they interact with other outcome measurements. The main function of the HAP is to empower Pac-12 clinicians by democratizing clinical data insights and data-driven decision-making in sports injury and illness. As the main contributor of clinical data to the HAP, certified ATs practicing at participating Pac-12 member schools can use the HAP analytic tools to inform patient outcomes and internal health care initiatives on their campus. This innovative feature also gives ATs access to actionable insight that can further solidify them as vital and valuable members of the clinical research community, both locally and throughout the Pac-12 Conference. Clinical applications of these tools include interactive dashboards that allow the user to explore aggregated data and drill down to areas of interest to inform monthly, end-of-year, and yearly comparative injury analyses. Aggregated PacTrac dashboards allow clinicians to identify trends in participation data, informing coaches and other stakeholders of emerging injury risk factors, including daily monitoring of return-to-play scenarios.

Another primary objective of the HAP is to supply deidentified clinical data to approved research groups in the Pac-12 Conference for secondary research use. The HAP project data are not accessible to the public but only to sports medicine clinicians and researchers associated with Pac-12 member schools. The HAP project data are refreshed quarterly to provide approved research groups access to current and relevant data. Since 2020, the HAP has serviced 33 formal research requests pertaining to a wide array of research questions directly affecting student-athlete health and well-being. The comprehensive application, data-use assurances, and dedicated personnel to service data requests, along with established relationships across the conference, allow relevant and high-impact data requests to be certified in a timely manner, minimizing the time to application of results. More specifically, in 2021, HAP data contributed to multiple descriptive epidemiologic studies[24,25,26] as well as to establishment of best-practice guidelines on COVID-19 testing and vaccination protocols by contributing HAP data to the Centers for Disease Control and Prevention to support optimal student-athlete quarantine periods.[13,27] The HAP infrastructure also sponsored efforts to compare the rate of concussions occurring during targeting versus nontargeting plays in American collegiate football, supporting concussion screening after plays in which targeting was involved.[28] Other examples of the HAP being used to improve student-athletes' health and wellness include investigations looking at the effect of overuse and non–time-loss injuries and validating biannual mental health screenings.[29]

Finally, the HAP uses innovative elements that other sports medicine organizations and conferences, as well as clinical settings, might treat as a blueprint when building their own comprehensive sports injury and illness analytics initiatives. First, the engagement of ATs in not only the collection of longitudinal point-of-care data but more importantly, in the development and oversight of the HAP, including the research portal, is a deliberate approach by the Pac-12 Conference to facilitate the type of collaborative relationships between clinicians and researchers recommended by the Athletic Training Research Agenda as necessary for conducting research that is relevant to ATs and other clinicians and most likely to positively affect student-athlete health and well-being.[30] Second, though the scale of the HAP may not be replicable by conferences with fewer resources or in secondary school settings, all or some of the breadth of data captured in the HAP could be easily captured with existing EHRs and used locally to inform patient care. Furthermore, the implementation of cost-effective cloud resources, software, and documentation workflows may allow organizations across many clinical settings to enhance their approach to clinical data analytics specific to their patient populations with minimal additional expense. By accommodating, and being designed to support both researchers and clinicians, the HAP provides a blueprint for the next generation of sports injury analytics systems to enhance the value proposition for ATs, who often serve as the primary data collectors for these systems. Therefore, whereas it may not be possible for all ATs or organizations to implement the full scope of the HAP as described, it is possible to identify which aspects of the HAP might provide the most value to clinical practice and use these in their own health analytics program.

REFERENCES

1. Student-athlete health and well-being initiative. Pac-12 Conference. Accessed December 2021. https://pac-12.com/student-athlete-health-well-being-initiative
2. Alles WF, Powell JW, Buckley W, Hunt EE. The National Athletic Injury/Illness Reporting System 3-year findings of high school and college football injuries. *J Orthop Sports Phys Ther*. 1979;1(2):103–108. doi:10.2519/jospt.1979.1.2.103
3. Kerr ZY, Dompier TP, Snook EM, et al. National Collegiate Athletic Association Injury Surveillance System: review of methods for 2004–2005 through 2013–2014 data collection. *J Athl Train*. 2014;49(4):552–560. doi:10.4085/1062-6050-49.3.58
4. Harmon KG, Dick R. The relationship of skill level to anterior cruciate ligament injury. *Clin J Sport Med*. 1998;8(4):260–265. doi:10.1097/00042752-199810000-00002
5. Kerr ZY, Dompier TP, Snook EM, et al. National Collegiate Athletic Association Injury Surveillance System: review of methods for 2004–2005 through 2013–2014 data collection. *J Athl Train*. 2014;49(4):552–560. doi:10.4085/1062-6050-49.3.58
6. Li RT, Kling SR, Salata MJ, Cupp SA, Sheehan J, Voos JE. Wearable performance devices in sports medicine. *Sports Health*. 2016;8(1):74–78. doi:10.1177/1941738115616917
7. Ekegren CL, Gabbe BJ, Finch CF. Sports injury surveillance systems: a review of methods and data quality. *Sports Med*. 2016;46(1):49–65. doi:10.1007/s40279-015-0410-z
8. Putukian M, D'Alonzo B, Campbell-McGovern C, Wiebe D. The Ivy League–Big Ten Epidemiology of Concussion study: a report on methods and first findings. *Am J Sports Med*. 2019;47(5):1236–1247. doi:10.1177/0363546519830100
9. Orchard J, Rae K, Brooks J, et al. Revision, uptake and coding issues related to the open access Orchard Sports Injury Classifica-

tion System (OSICS) versions 8, 9 and 10.1. *Open Access J Sports Med*. 2010;1:207–214. doi:10.2147/OAJSM.S7715
10. Rae K, Orchard J. The Orchard Sports Injury Classification System (OSICS) version 10. *Clin J Sport Med*. 2007;17(3):201–204. doi:10.1097/JSM.0b013e318059b536
11. Bohr AD, Aukerman DF, Harmon KG, et al. Pac-12 CARE-Affiliated Program: structure, methods and initial results. *BMJ Open Sport Exerc Med*. 2021;7(2):e001055. doi:10.1136/bmjsem-2021-001055
12. Norcross MF, Johnson ST, Bovbjerg VE. Athletic exposure methodology influences injury risk estimates in collegiate sports [abstract]. *J Athl Train*. 2018;53(suppl 6):S58. doi:10.4085/1062-6050-53.6s.s1
13. Harmon K, de St Maurice AM, Brady AC, et al. Surveillance testing for SARS-COV-2 infection in an asymptomatic athlete population: a prospective cohort study with 123 362 tests and 23 463 paired RT-PCR/antigen samples. *BMJ Open Sport Exerc Med*. 2021;7(2):e001137. doi:10.1136/bmjsem-2021-001137
14. Broglio S, McCrea M, McAllister T, et al. A national study on the effects of concussion in collegiate athletes and US military service academy members: the NCAA–DoD Concussion Assessment, Research and Education (CARE) Consortium structure and methods. *Sports Med*. 2017;47(7):1437–1451. doi:10.1007/s40279-017-0707-1
15. Mental Health Coordinating Unit. Pac-12 student-athlete health & well-being grants program. Accessed December 2021. https://pac-12.com/mentalhealth#coordinatingunit
16. Brooks JHM, Fuller CW. The influence of methodological issues on the results and conclusions from epidemiological studies of sports injuries: illustrative examples. *Sports Med*. 2006;36(6):459–472. doi:10.2165/00007256-200636060-00001
17. Stovitz SD, Shrier I. Injury rates in team sport events: tackling challenges in assessing exposure time. *Br J Sports Med*. 2012;46(14):960–963. doi:10.1136/bjsports-2011-090693
18. NCAA 2021–22 Division I manual. National Collegiate Athletic Association. Accessed December 2021. web3.ncaa.org/lsdbi/reports/getReport/90008
19. Ward MJ, Self WH, Froehle CM. Effects of common data errors in electronic health records on emergency department operational performance metrics: a Monte Carlo simulation. *Acad Emerg Med*. 2015;22(9):1085–1092. doi:10.1111/acem.12743
20. Eekhout I, de Vet HCW, Twisk JWR, et al. Missing data in a multi-item instrument were best handled by multiple imputation at the item score level. *J Clin Epidemiol*. 2014;67(3):335–342. doi:10.1016/j.jclinepi.2013.09.009
21. Mack CD, Meisel P, Herzog MM, et al. The establishment and refinement of the National Basketball Association Player Injury and Illness Database. *J Athl Train*. 2019;54(5):466–471. doi:10.4085/1062-6050-18-19
22. Langkamp DL, Lehman A, Lemeshow S. Techniques for handling missing data in secondary analyses of large surveys. *Acad Pediatr*. 2010;10(3):205–210. doi:10.1016/j.acap.2010.01.005
23. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst*. 2014;2:3. doi:10.1186/2047-2501-2-3
24. Schroeder GG, McClintick DJ, Trikha R, Kremen TJ Jr. Injuries affecting intercollegiate water polo athletes: a descriptive epidemiologic study. *Orthop J Sports Med*. 2022;10(7):23259671221110208. doi:10.1177/23259671221110208
25. Trikha R, Schroeder GG, Greig DE, Kremen TJ Jr. Characterizing health events and return to sport in collegiate swimmers. *Orthop J Sports Med*. 2022;10(4):23259671221083588. doi:10.1177/23259671221083588
26. Brown C, Bovbjerg V, Soucy M, Fredericson M, Simon J. Acute and overuse, time-loss and non-time-loss injuries in collegiate student-athletes [abstract]. *Med Sci Sports Exerc*. 2021;53(suppl 8):205. doi:10.1249/01.mss.0000761448.58818.c1
27. Atherstone C, Peterson ML, Malone M, et al. Time from start of quarantine to SARS-CoV-2 positive test among quarantined college and university athletes—17 states, June–October 2020. *MMWR Morb Mortal Wkly Rep*. 2021;70(1):7–11. doi:http://dx.doi.org/10.15585/mmwr.mm7001a2
28. Aukerman DF, Bohr AD, Poddar SK, et al. Risk of concussion after a targeting foul in collegiate American football. *Orthop J Sports Med*. 2022;10(2):23259671221074656. doi:10.1177/23259671221074656
29. Student-Athlete Health and Well-Being Initiative. Pac-12 Conference. Accessed December 2021. https://pac-12.com/pac-12-research-grants-program
30. Eberman LE, Walker SE, Floyd RT, et al. The Prioritized Research Agenda for the Athletic Training Profession: a report from the Strategic Alliance Research Agenda Task Force. *J Athl Train*. 2019;54(3):237–244. doi:10.4085/1062-6050-374-18

Address correspondence to Kevin C. Robell, MA, ATC, ARA Medical Enterprises Inc, 2305 Historic Decatur Road, Suite 100, San Diego, CA 92106. Address email to kevin@preventicx.com.