Abstract: Updates to industry consensus standards for ECG equipment is a work-in-progress by the ISO/IEC Joint Work Group 22. This work will result in an overhaul of existing industry standards that apply to ECG electromedical equipment and will result in a new single international industry, namely 80601-2-86. The new standard will be entitled “80601, Part 2-86: Particular requirements for the basic safety and essential performance of electrocardiographs, including diagnostic equipment, monitoring equipment, ambulatory equipment, electrodes, cables, and leadwires”. This paper will provide a high-level overview of the work in progress and, in particular, will describe the impact it will have on requirements and testing methods for computerized ECG interpretation algorithms. The conclusion of this work is that manufacturers should continue working with clinical ECG experts to make clinically meaningful improvements to automated ECG interpretation, and the clinical validation of ECG analysis algorithms should be disclosed to guide appropriate clinical use. More cooperation is needed between industry, clinical ECG experts and regulatory agencies to develop new data sets that can be made available for use by industry standards for algorithm performance evaluation.

Keywords: ECG equipment; computerized electrocardiograph; ECG analysis algorithms; computerized ECG interpretation

1. Introduction

Industry standards are published for particular types of electromedical equipment by the International Electrotechnical Commission (IEC) and the International Organization for Standardization (ISO). These industry standards are updated on a regular basis by ISO/IEC workgroups. Work that is in progress by the Joint Workgroup 22 (JWG22) under the ISO/IEC 62D Electromedical Equipment Subcommittee will result in the publication of a new standard for ECG devices and systems with the designation of ISO/IEC 80601-2-86, which will be entitled “80601, Part 2-86: Particular requirements for the basic safety and essential performance of electrocardiographs, including diagnostic equipment, monitoring equipment, ambulatory equipment, electrodes, cables, and leadwires” [1]. JWG22 is a joint workgroup formed between the maintenance team that oversees the ECG particular standards and liaisons from other standard workgroups. This new standard is currently in draft form and constitutes a significant overhaul of current ECG equipment standards and, in effect, combines the three current ECG particular standards published by the IEC for diagnostic electrocardiographs [2], ECG patient monitors [3], and ambulatory ECG equipment [4]. The standard will additionally incorporate three ECG-related standards published by the Association for the Advancement of Medical Instrumentation (AAMI), which is the national standard development organization in the United States for health technology. The three additional AAMI standards will have safety and performance requirements for disposable electrodes (AAMI EC12 [5]), ECG cables and leadwires (AAMI EC53 [6]), and arrhythmia analysis performance reporting (AAMI EC57 [7]). Finally, 80601-2-86 will restore requirements that were omitted from a previously deprecated IEC diagnostic ECG particular standard that addressed the performance of computerized ECG analysis [8]. The enormous effort required for the development of the new 80601-2-86 standard represents...
a formidable task with far reaching implications, such that a comprehensive discussion of changes is beyond the scope of this paper. Therefore, the intention of this paper is narrowed in focus to give the reader a cursory level of understanding of the work in progress and a more detailed discussion about the impact it will have, specifically on performance requirements for computerized diagnostic ECG analysis algorithms, which is also commonly referred to by other terms, such as automated ECG interpretation or computerized ECG interpretation.

2. Background

Industry standards for ECG equipment have existed since the 1980s and were historically developed by separate workgroups for different types of ECG equipment and across different standard organizations, such as the IEC or AAMI. This has resulted in a complex landscape of industry standards, which are recognized at different levels for compliance by different regulatory agencies. Moreover, specific standards were developed for different types of ECG devices (namely diagnostic electrocardiographs, ambulatory ECG equipment, and ECG patient monitors). It is now common that contemporary ECG equipment does not clearly fit into any of these historical definitions of specific types of ECG devices, which has led to confusion and resulted in the inconsistent use of these standards. This is a challenge for manufacturers, clinicians, regulatory agencies, and testing facilities alike. Table 1 contains the list of the different historical standards that have been applied to ECG equipment. Despite the efforts that have been made over the years by the AAMI and IEC organizations to harmonize redundant ECG standards [9], challenges continue to exist because the current standards were developed decades ago and very few changes have been made over the years to requirements or conformance testing methods to update the standards with advancements made in either ECG technology or the clinical use of the ECG. Consequently, there is a tremendous opportunity recognized by the standard developers in JWG22 to accomplish three primary goals: combine current standards into a single standard that will encompass all types of ECG medical equipment, harmonize requirements and testing methods, and make updates to bring requirements and conformance testing methods into alignment with the current state of the art for ECG equipment and clinical use of ECG. These goals, albeit a monumental effort, have been adopted for current work in progress by JWG22 and will result in the new single ISO/IEC 80601-2-86 ECG equipment standard.

The first committee draft of 80601-2-86 [1] has been published and the second committee draft is in progress at the time of preparing this paper. Draft 80601-2-86 is organized as a set of general basic safety and performance requirements that will apply to all ECG devices, together with additional clauses that retain specific requirements for ECG equipment included within the definitions of diagnostic electrocardiographs, ECG patient monitors, and ambulatory ECG equipment. The definitions for these different types of ECG equipment are based on the intended use for the ECG equipment claimed by the manufacturer. A fundamental goal of the 80601-2-86 standard is to provide updated definitions for these specific types of ECG equipment linked to intended use. Explanation of these intended use definitions have been updated with guidance intended to make it easier to understand how current types of devices as well as emerging novel devices fit into defined categories of ECG equipment. These updates should provide better clarity and understanding regarding how the new standard should be applied to existing types of ECG devices as well as future innovations.
Table 1. List of historical industry standards that apply to ECG equipment.

<table>
<thead>
<tr>
<th>Standard Designation</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAMI EC11 [10]</td>
<td>Diagnostic electrocardiographic devices</td>
</tr>
<tr>
<td>AAMI EC13 [11]</td>
<td>Cardiac monitors, heart rate meters, and alarms</td>
</tr>
<tr>
<td>AAMI EC38 [12]</td>
<td>Ambulatory electrocardiographs</td>
</tr>
<tr>
<td>AAMI EC53 [6]</td>
<td>ECG cables and leadwires</td>
</tr>
<tr>
<td>AAMI EC57 [7]</td>
<td>Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms</td>
</tr>
<tr>
<td>IEC 60601-2-51 [8]</td>
<td>Medical electrical equipment—Part 2-51: Particular requirements for the basic safety and essential performance of recording and analyzing single channel and multichannel electrocardiographs</td>
</tr>
</tbody>
</table>

3. Impact of 801601-2-86 on Automated ECG Interpretation

One of the key aspects impacted by this work is the update to requirements for computerized analysis of ECG signals. Great efforts have been made in this new standard to combine the different requirements and testing methods and clarify how they should be applied to different types of ECG analysis algorithms in a single standard. There are currently different sets of requirements, testing methods, and test data sets defined in existing standards [2–4,7]. While it may seem to be a simple task to combine the algorithm testing requirements, methods and data sets from the existing standards, and to add clarifications and rationale, 80601-2-86 addresses a long-standing challenge to understand the scope and purpose of the different algorithm testing requirements, as well as how to apply them across different types of ECG equipment. Historical ECG device standards have each had clauses that apply to ECG analysis algorithms [2–4,7,8,10–13] along with corresponding definitions of the types of ECG devices to which they apply. Unfortunately, the definitions focus more on the type of device containing the algorithm rather than the intended clinical use of the algorithm. Moreover, there was no guidance for manufacturers regarding how to apply these standards to ECG analysis algorithms that were contained in ECG equipment that did not meet these specific ECG device definitions.

When the new 80601-2-86 standard is introduced, ECG algorithm testing requirements, testing methods and data sets will be applied based on the intended use of the algorithm and not just the type of ECG device, which contains the algorithm. The requirements in the existing draft are structured with two different clauses, namely 201.12.4.1 Algorithm testing for Diagnostic 12 Lead and 201.12.4.2 requirements for testing computerized arrhythmia analysis algorithms [1]. Requirements in each of these clauses are applied to the specific types of ECG equipment for which they were originally defined in historical standards. In addition, these requirements are also applied to other computerized ECG analysis algorithms based on the intended use of the computerized ECG analysis output rather than the definition of the equipment itself.

In general, the diagnostic 12 lead algorithm runs on a static recording of an ECG snapshot to generate measurements, rhythm interpretation and may include interpretive statements for conduction, and morphologic patterns [1] for abnormalities that may include a wide range of diseases such as hypertrophic disease, ischemic disease, acute myocardial infarction, primary and secondary repolarization abnormalities, etc. Algorithms that are intended to provide a diagnostic 12 lead ECG interpretation using data that are derived from a non-standard reduced lead set, such as the EASI system [14] or from a reduced
precordial lead [15], also meet the description of a diagnostic 12 lead algorithm, even though the devices which may contain these algorithms do not meet the definition of a diagnostic electrocardiograph (“DIAGNOSTIC ECG ME EQUIPMENT” [1]) in 80601-2-86.

In contrast, arrhythmia analysis algorithms are intended to analyze data in a more continuous nature and may analyze long term data, such as those from Holter or ECG patch devices, may analyze continuous and/or real time data, such as those from ECG patient monitors, or may analyze short term ECG data, such as those from ECG event recorders or mobile cardiac telemetry (MCT) type devices. The intended purpose of arrhythmia algorithms is to detect and classify QRS complexes and detect arrhythmic events [1]. Arrhythmia analysis algorithms may also perform ECG measurements for the purpose of trending measurement or detecting events, such as ischemic episodes.

There is some overlap between the outputs of these two types of ECG analysis algorithms, but they have different intended uses, and, therefore, the requirements, testing methods, and testing data sets are different for each of these two types of algorithms. The following discussion will focus on the impact of 80601-2-86 on performance testing for diagnostic 12 lead ECG analysis algorithms, which are also referred to by other descriptions, such as “automated ECG interpretation”. The statistical metrics, limitations of testing and underlying principles for automated ECG interpretation also apply to arrhythmia analysis algorithms as well but will not be discussed in this paper.

Most current diagnostic electrocardiographs now have the ability for computer automated ECG interpretation and, by 2006, it was estimated that 100 million ECGs were being interpreted by computerized algorithms in the United States and a similar number in Europe and in the rest of the world [16]. The performance of these computerized ECG analyses has reached a point where the algorithms can make routine ECG measurements accurately and provide useful clinical benefits, yet also having well studied limitations when compared to humans over reading [17]. Because of the widespread use of computer-automated ECG interpretation and the impact it can have on clinical decision making, it is imperative to include the performance-testing requirements that provide as comprehensive a characterization of the algorithm performance as possible. This goal has been a cornerstone part of industry standards for electrocardiographic equipment and is maintained in the 80601-2-86 standard. It is based on requirements and testing methods developed by the Common Standards for Quantitative Electrocardiography (CSE) project [18,19] and the European Conformance Testing Services (CTS) project [20].

Current requirements for testing diagnostic ECG interpretation algorithms only require testing the accuracy of amplitude measurements and interval measurements on CTS and CSE data using calibration, analytic and biologic waveforms [2]. CTS analytic and calibration ECG waveforms are both simulated ECG-like waveforms with a range of characteristics. Calibration ECG waveforms are artificial in nature and designed to test both the hardware response of a device, as well as the automated ECG measurement performance used for automatic diagnostic ECG interpretation programs. Analytic ECG waveforms are more physiologically realistic in nature and designed to measure the accuracy of 12 lead diagnostic ECG interpretative programs in the detection and measurement of the ECG features. Figure 1 shows examples of CTS calibration and analytic waveforms. Biologic waveforms consist of a small set of actual physiologic ECG recordings that have been annotated by human over readers.
Figure 1. Examples of simulated waveforms from the CTS database used for algorithm performance testing in current ECG standards [2]. (a) Example of a calibration ECG waveform with ECG waves and interval nomenclature indicated; (b) example of an analytic ECG waveform with ECG waves and interval nomenclature indicated.

Diagnostic statements for automated ECG interpretation algorithms are the fundamental output, and the accuracy of these statements should be well characterized by algorithm testing, including both ECG contour and ECG rhythm diagnostic statements. The methods for measuring accuracy have been consistently and well defined, and previously included in ECG standards [8]. However, the current industry ECG standards have omitted these historical requirements to test the accuracy of diagnostic statements [2], which is a gap that is being addressed in 80601-2-86. This testing is particularly important with the emergence of new types of algorithms, such as machine learning, for which little guidance is available regarding the validation of clinical accuracy. The challenge with making even further
improvements to this situation is that advancement requires better test datasets and, at this time, there are no new available databases with appropriate types of data including properly adjudicated reference annotations, and which are publicly accessible for inclusion in an industry standard. Consequently, improvements that are being made in 80601-2-86 are limited in nature.

The content of 80601-2-86 combines and harmonizes the safety and performance requirements from historical ECG standards. This includes requirements that address the technical aspects of signal acquisition and signal conditioning to ensure that ECG equipment will operate safely and acquire signals that are appropriate for the intended use of the equipment, which may include both human interpretation and/or the computer analysis of the ECG [21]. These technical requirements address necessary performance specifications to ensure that resulting signals are appropriate for their intended use and include specifications, such as filtering, bandwidth, common mode rejection, and system noise [22]. It is beyond the scope of this paper to discuss the effects of inadequate signal acquisition and signal conditioning on the effects of computerized ECG interpretation. It is assumed that, if ECG equipment is compliant with conformance testing for ECG signal acquisition, then the output ECG signals will be appropriate for both human and computerized interpretation, based on the intended use claimed by the manufacturer.

There are strong data to support the proposition that computerized ECG interpretation programs provide an important clinical adjunct to the physician that may even enhance physician overreading [23,24], but it is also clearly understood that the outputs of all computerized ECG interpretation algorithms have limitations [25] and require physician overread [21,26]. The historical requirements, testing methods and testing data sets have changed little over the years. Methods for measuring automated ECG interpretation have been consistently applied over the years by current [2] and past [8] industry standards. However, the data used for testing can heavily influence the measurement of accuracy and, at this time, there are no new additional databases that are appropriate for use as an industry standard, although some new efforts are ongoing [27]. Consequently, little progress has been made in improving the current quality of performance testing for algorithms in 80601-2-86. The work required to create better reference data sets for algorithm testing is particularly daunting and the improvements that can be made to the current performance testing are limited until better data sets are available for use within the context of an industry standard.

Developers will continue to improve the accuracy of automated ECG interpretation programs and individual manufacturers will continue to validate algorithm performance with private data sets. Furthermore, the emerging use of machine learning and artificial intelligence algorithms for ECG interpretation will add new complexities to the problem of understanding and characterizing algorithm safety and performance. This is also challenging regulatory agencies to expand their considerations for algorithm development and validation to address these new complexities [28]. Nevertheless, because of the ubiquitous presence of automated ECG interpretation software and because of the impact it can have on clinical diagnosis and decision making, it remains critical for the performance evaluation of algorithms to be a compulsory element of industry standards for ECG equipment.

It has been debated through the years as to the value and validity of some of testing requirements that have been included in ECG equipment standards. In fact, when the IEC 60601-2-51 standard was combined with the IEC 60601-2-25 standard, two areas of algorithm testing were not included in the update, namely (1) the testing requirements that pertain to the evaluation of diagnostic ECG measurements in the presence of noise and (2) reporting for interpretive 12 lead diagnostic statements [2]. At the time when these two standards were combined, the consensus of the workgroup was that they had limited value. However, the consensus of the JWG22 workgroup has changed based on constituency feedback and now acknowledges that these deprecated algorithm-testing requirements are important elements of computerized ECG interpretation and should be mandatory to improve algorithm-testing requirements. Reviews of the use of the CSE and
CTS data sets in 80601-2-86 clearly indicate that these test data sets have limitations [29], and consequently manufacturers often use proprietary data sets for validating clinical performance of automatic ECG interpretation. Manufacturers must be cognizant that measuring diagnostic accuracy depends on the quality of the data composition and should use data that are representative of the intended clinical environment and consider that the predictive merits of performance evaluation must be examined in relation to sample size, patient populations and the prevalence values for each diagnostic category.

In general, it is expected that each diagnostic category should be validated by an adequate number of clinical cases, and the use of enriched datasets may be acceptable if supplemented by supporting analysis. Noise is a common occurrence in clinical environments and the variation of errors will increase with degraded signal quality [30] and the impact of noise on the measurement accuracy of diagnostic ECG measurement algorithms is an important characteristic to evaluate [31]. In general, a manufacturer should assess whether the databases and methods defined by the standard are fully representative of the device under test as well as its use and determine when deviations or additional testing may be needed (e.g., additional device-specific data, numerical transformation, additional noise patterns, etc.). While methods included in 80601-2-86 were originally developed for tradition rule-based algorithms (i.e., those that implement classification rules based on clinical consensus), the general concepts for testing and reporting may also be applied to algorithms based on machine learning/artificial intelligence, although larger testing datasets and additional analysis may be needed to ensure a robust validation.

It is important to note that 80601-2-86 does not specify pass–fail criteria for automated ECG analysis performance. This obviously does not mean that any performance is acceptable; instead, it is a recognition that the performance of an automated ECG analysis algorithm should be evaluated in the context of the device’s intended use, to ensure that the device performs sufficiently well in clinical practice.

4. Discussion

At the time of preparing this paper, the first committee draft of 80601-2-86 had been published and circulated for comments by national standard organizations members of the IEC JWG22. The second committee draft is in preparation for circulation to the national committees for a second call for comments. The current state of 80601-2-86 combines several existing standards that apply to ECG equipment into a single standard that will include all ECG equipment within its scope and will also contain specific requirements for particular types of ECG equipment based on intended use claimed by the manufacturer. This will include requirements and conformance testing methods for computerized ECG analysis algorithms, which are defined in two broad categories, namely diagnostic 12 lead ECG interpretative algorithms and arrhythmia analysis algorithms. The quantification of performance and testing data sets have been in existence for decades. The goal of the new 80601-2-86 standard is to update the rationales and guidance contained in the informative annexes in such a way that it is more clearly understood how to apply the standard to the range of contemporary computerized ECG analysis algorithms based on the intended use of the ECG equipment in which they are used.

In particular, the requirements for measurement and analysis algorithms for diagnostic ECG interpretation restore some historical performance testing requirements and conformance testing methods that had been previously deprecated from current standards. Although the limitations of the conformance testing data sets have been well recognized and published, they still provide the only method of uniformly and consistently benchmarking algorithm performance. This is especially important because of the ubiquitous use of automated ECG interpretation by the clinical community and the important influence it can have on physician over reading.

Furthermore, the profound influence that automated ECG interpretation programs can have on physician ECG interpretation and clinical decision making has been well published by experts in electrocardiography and the importance of developing and evaluating
the performance of these algorithms with scientific rigor is critical to ensuring that the appropriate use of computerized ECG interpretation programs is well understood and benefits patient care.

While the 80601-2-86 standard applies to the vast majority of ECG devices, the requirements for automated ECG analysis and interpretation are mostly relevant for traditional device types (e.g., rule-based analysis of resting 12-lead ECG and traditional Holter ECGs). However, the same concepts can be applied to novel technologies (e.g., machine learning/AI-based algorithms, non-standard lead technology/lead configuration) by using additional datasets relevant for the device’s intended use. Manufacturers should pay particular attention to factors that impact the quality and appearance of data sets for both algorithm development and testing, in particular, establishing appropriate sample sizes, patient population representation, and disease prevalence/representation to accurately reflect the clinical environment and intended use for which the algorithm is designed.

5. Summary

The introduction of 80601-2-86 is a significant overhaul of existing industry standards and will result in a single international standard that can be applied to all ECG equipment. The goals are to combine, update, and harmonize the safety and performance requirements from the multiple existing industry standards so that the new standard can be applied to all types of ECG and be appropriate for current ECG technology and clinical use.

Although the CSE and CTS test data sets have well known limitations, no other data sets have been accepted for inclusion in 80601-2-86. Manufacturers should continue to work together with clinical ECG experts to continue clinically meaningful improvements to computerized ECG analysis and should disclose the clinical validation of algorithm improvements to guide appropriate clinical use. More cooperation is needed between industry, clinical ECG experts and regulatory agencies to develop new data sets that can be made available for use by industry standards for algorithm performance evaluation.

Author Contributions: All authors have been directly involved with work to draft the ISO/IEC 80601-2-86 particular standard as active members of JWG22. J.-J.S. is the current convenor for JWG22. B.Y. also serves as a Co-Chair of the AAMI ECG Committee. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References


