

R. Bousseljot
D. Kreiseler

Ergebnisse der EKG-Interpretation mittels Signalmustererkennung

Eingegangen: 12. Oktober 2000
Akzeptiert: 15. November 2000

Results of ECG interpretation by means of signal waveform recognition

Summary This contribution presents the development of the pattern recognition procedure for ECG signal evaluation. In contrast to the methods already available, the procedure is not based on an ever deeper splitting of the signal into individual attributes but uses procedures of signal pattern recognition in conjunction with ECG databases. With the aid of an ECG signal pattern database with 12000 ECGs, sensitivities and specificities for the main groups of findings are presented and discussed for 8500 examinations. Conclusions for the further development of the procedure are drawn.

Key words ECG evaluation – waveform recognition – multi-channel signal analysis – cross correlation – computer diagnostics – ECG signal pattern database

Zusammenfassung Im vorliegenden Beitrag wird die Weiterentwicklung des Mustererkennungsverfahrens zur EKG-Signalauswertung vorgestellt. Dieses Verfahren beruht im Gegensatz zu bisherigen Methoden nicht auf einer immer tiefer gehenden Zerlegung des EKG-Signals in Einzelmerkmale, sondern nutzt Verfahren der Signalmustererkennung in Verbindung mit EKG-Datenbanken. Anhand einer EKG-Signalmusterdatenbank mit 12000 EKG werden für 8500 Untersuchungen Sensitivitäten und Spezifitäten für die Hauptbefundgruppen vorgestellt und diskutiert. Schlussfolgerungen für die Weiterentwicklung des Verfahrens werden abgeleitet.

Schlüsselwörter EKG-Auswertung – Wellenformerkennung – multikanalige Signalanalyse – Kreuzkorrelation – Computerdiagnostik – EKG-Signalmusterdatenbank

R. Bousseljot (✉) · D. Kreiseler
Physikalisch-Technische Bundesanstalt
Labor 8.21 Bioelektrizität
und Biomagnetismus
Abbestraße 2–12
10587 Berlin

Einleitung

Zur rechnergestützten Interpretation von EKG werden Verfahren eingesetzt, die durch Vermessung der Wellenformen des EKG Merkmale aus den Signalen als Parameter extrahieren (1–3). Mit Hilfe dieser

Merkmalsätze werden durch regelbasierte oder statistische Klassifikatoren oder durch Einsatz neuronaler Netze (4, 5) Befundvorschläge abgeleitet.

In (6–8) wurde ein neues Verfahren zur Gewinnung diagnostischer Informationen aus dem EKG-Signal vorgestellt, das nicht mehr auf der

regelbasierten Vermessung und Extraktion von Einzelmerkmalen aus dem EKG beruht. Durch Vergleich der Signalmuster eines 12-Kanal-EKG unbekannter Diagnose (Test-EKG) mit Signalmustern einer EKG-Datenbank (Referenz-EKG) bekannter Diagnose werden diejenigen EKG aus der Datenbank ge-

sucht, die in ihren Wellenformen die größte Ähnlichkeit zum Test-EKG besitzen. Aus den Diagnosen der selektierten Datenbank-EKG wird auf die Diagnose des unbekannten EKG geschlossen.

Die Funktion des Verfahrens wurde 1998 an Hand einer EKG-Datenbank mit ca. 560 EKG vorrangig von Patienten mit Myokardinfarkt sowie gesunden Probanden getestet (6–8). Damit konnte die grundlegende Funktionsfähigkeit des Verfahrens für die Erkennung von Patienten mit Myokardinfarkten sowie von Probanden nachgewiesen werden (6). Aufgrund des unzureichenden Umfangs der Datenbasis war ein Nachweis der Anwendbarkeit der Methode auf die Befundung von EKG anderer Diagnosegruppen nicht möglich.

Inzwischen konnte die EKG-Signalmusterdatenbank auf ca. 12 000 EKG erweitert werden. Nachfolgend werden die in Verbindung mit dem Aufbau dieser Datenbank erfolgten Weiterentwicklungen des Verfahrens und die erreichten Ergebnisse vorgestellt.

Kurzbeschreibung des Interpretationsverfahrens

Ausgangsüberlegung für das in (6, 7) beschriebene Verfahren ist die Annahme, dass zwei EKG, die in ihren jeweils gleichen medizinischen Ableitungen gleiche oder sehr ähnliche Signalmuster besitzen, auch die gleichen oder sehr ähnlichen kardiologischen Diagnosen haben. Wenn diese Hypothese richtig ist, sollte es möglich sein, durch einen Mustervergleich der Ableitungen eines unbekannten Test-EKG mit den Ableitungen in einer Datenbank erfasster Referenz-EKG mit bekannter Diagnose auf die Diagnose des unbekannten Test-EKG zu schließen. Dazu werden in einer EKG-Signaldatenbank EKG gesucht, deren Ableitungen die größte Übereinstimmung mit dem Signalmuster der je-

weiligen Ableitungen des Test-EKG besitzen. Dies erfolgt durch die Berechnung der Korrelationsfunktion mit Hilfe der Gleichung 1 (7, 8).

Übereinstimmung der Signalmuster der Ableitungen des Test-EKG mit den Ableitungen des Referenz-EKG darstellen.

$$K = \frac{\sum_{n=1}^N x_n y_n - \frac{1}{N} \sum_{n=1}^N x_n \sum_{n=1}^N y_n}{\sqrt{\left[\sum_{n=1}^N x_n^2 - \frac{1}{N} \left(\sum_{n=1}^N x_n \right)^2 \right] \left[\sum_{n=1}^N y_n^2 - \frac{1}{N} \left(\sum_{n=1}^N y_n \right)^2 \right]}} \quad (1)$$

Im Verlauf dieser Berechnungen wird jede Ableitung des Test-EKG mit den entsprechenden Ableitungen der Referenz-EKG der Datenbank verglichen. Um Rechenzeit zu sparen, wird für das Test-EKG ein repräsentativer Signalausschnitt verwendet, der mindestens einen vollständigen Herzschlag enthält.

Die Bilder 1 und 2 zeigen diese Korrelationsfunktion für eine gut (Bild 1) sowie eine schlecht (Bild 2) übereinstimmende EKG-Ableitung. Für die weitere Auswertung werden die Maxima der Korrelationsfunktion verwendet, die ein Maß für die

Die so erhaltenen Ergebnisse für die einzelnen Ableitungen werden mit einem multivariaten Klassifizierungsverfahren für jedes Referenz-EKG zusammengefasst (7, 8). Im Ergebnis dieser Berechnungen lassen sich die Referenz-EKG nach dem Grad der Übereinstimmung zum Test-EKG sortieren. Den am besten mit dem zu untersuchenden Test-EKG übereinstimmenden Referenz-EKG können nun weitere Informationen aus der Datenbank, beispielsweise Befundaussagen, zugeordnet werden. Um auch Mehrfachbefunde zu berücksichtigen, werden jedem Referenz-EKG bis zu fünf einzelne Aussagen zugeordnet.

Die Befundaussagen werden entsprechend der Wahrscheinlichkeit ihres Zutreffens mit einem Faktor von 0–100 gewichtet (Konfidenzniveau). Für die Auswertung wird eine vorbestimmte Anzahl von Referenz-EKG mit der besten Signalmusterübereinstimmung mit dem Test-EKG herangezogen (Auswertetiefe). Für vorliegende Untersuchungen wurde die Auswertetiefe 50 gewählt. Für diese Referenz-EKG werden die entsprechenden Befundaussagen unter Berücksichtigung ihrer Wichtung summiert und in einem Balkendiagramm (Bild 3) grafisch dargestellt. Auf der X-Achse des Diagramms sind die ausgewerteten Hauptbefundgruppen und auf der Y-Achse deren prozentuale Aufteilung dargestellt. Als Hauptbefundgruppen (siehe Legende) wurden folgende Befunde gewählt.

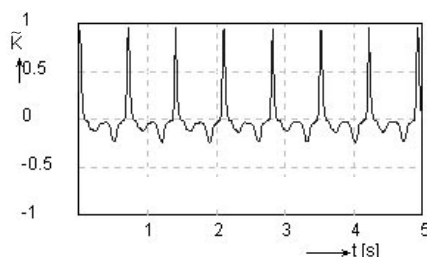


Abb. 1 Beispiel einer Korrelationsfunktion mit deutlicher Ausprägung lokaler Maxima

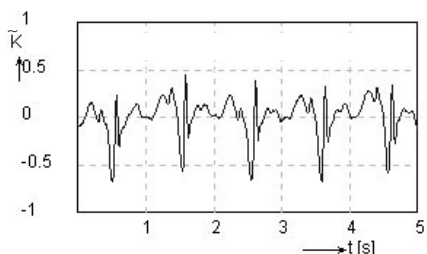


Abb. 2 Beispiel einer Korrelationsfunktion ohne lokale Maxima (keine Ähnlichkeit der Muster)

Tab. 1 Bedeutung der Hauptbefunde der Legende in Bild 3

Kurzbezeichnung Legende	Befundgruppe
Unbekannt	Unspezifischer Befund
NORM	Normales EKG
POSNL	Möglicherweise normales EKG
HYP	Hypertrophie
MI	Myokardinfarkt
Block	Block
ISC	Ischämie

Unter den Diagrammbalken ist die absolute Anzahl der zu den einzelnen Befundgruppen gefundenen Befundaussagen (einschließlich der Unterbefunde) angegeben. Die Prozentzahl über den Balken berücksichtigt die Wichtung der jeweiligen Befundaussagen. Die summierten Maßzahlen aller entsprechend der Auswertetiefe 50 berücksichtigten EKG ergeben 100%. Damit lässt sich für jede Befundgruppe eine prozentuale Wahrscheinlichkeit ihres Zutreffens angeben.

Bild 3 zeigt als Beispiel die Klassifikation nach Hauptbefundgruppen. Wie dem Bild zu entnehmen ist, wurde unter den Befundaussagen der ähnlichsten 50 EKG insgesamt 59-mal ein Befund „Block“ und 8-mal der Befund „Myokardinfarkt“ gefunden. Bei dieser Betrachtung werden alle zum jeweiligen Hauptbefund gehörenden Unterbefunde mit erfasst. Das bedeutet, 84,90% aller gewichteten Befundaussagen der 50 EKG mit den ähnlichsten Signalmustern entsprechen dem Befund „Block“. In gleicher Weise kann in einem weiteren Schritt der Befundhinweis „Block“ weiter differenziert werden. Bild 4 zeigt die weiterführende Analyse des Beispiels Bild 3. Von den 59 Befundaussagen „Block“ entfallen 45 (74,70%) auf die Untergruppe „Kompletter Rechtsschenkelblock“ (CRBBB), 8 auf die Untergruppe „links anteriorer Hemiblock“ (LAFB) sowie 4 (16,10%)

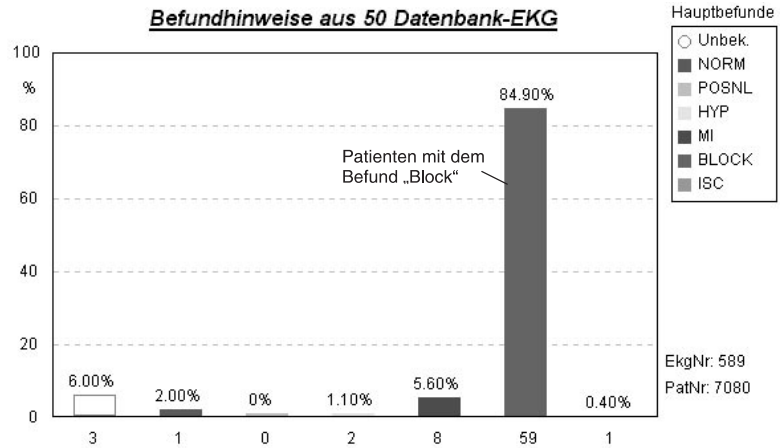


Abb. 3 Beispiel für mehrdimensionale gewichtete Zuordnung der Klassifikationsergebnisse zu kardiologischen Hauptbefundgruppen

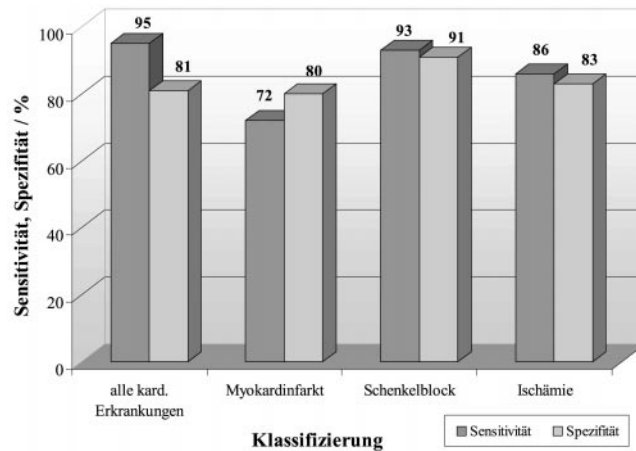


Abb. 4 Sensitivitäten und Spezifitäten der Interpretationsergebnisse bei kardiologischen Hauptbefundgruppen (höchstes Konfidenzniveau)

auf die Untergruppe „Unspezifischer Block“. Damit kann man für das zu untersuchende EKG mit sehr hoher Wahrscheinlichkeit auf das Vorliegen eines kompletten Rechtsschenkelblocks schließen.

Die in den Bildern 3 und 4 beispielhaft gezeigten Zuordnungen der Klassifikationsergebnisse wurden für folgende Merkmals- bzw. Merkmalsuntergruppen realisiert:

- ohne Klassenzuordnung
- Kardiologische Hauptbefunde
- Infarktstadium
- Hypertrophien
- Normal/Pathologisch
- Infarktlage
- Blocktyp

Die vorliegende Programmlösung erlaubt das Hinzufügen weiterer Merkmale, soweit entsprechende Datenbankinformationen vorhanden und die zu berücksichtigenden Merkmale im Signalmuster der EKG enthalten sind. Zur Festlegung von weiteren Merkmalen oder Merkmalskombinationen sind weitere Untersuchungen notwendig.

Datenbasis des Verfahrens

Grundlage des Verfahrens ist eine umfangreiche, möglichst gut validierte EKG-Signalmuster-Datenbank.

Dabei sind für die Analyseergebnisse die Qualität und der Wahrscheinlichkeitsgrad der Befundaussagen der Datenbank-EKG von großer Bedeutung. Die Genauigkeit der Befundhinweise des Mustererkennungsverfahrens kann nicht besser sein, als die summarische Genauigkeit der in der Datenbasis enthaltenen Befundangaben.

Ebenso wichtig ist die Anzahl der Referenz-EKG der Datenbank, die das biologische Spektrum der kardiologischen Erkrankungen weitgehend widerspiegeln muss. EKG mit z.B. seltenen kardiologischen Erkrankungen oder Kombinationen davon, die in der Datenbasis nicht enthalten sind, können durch das Mustererkennungsverfahren nicht selektiert und bewertet werden.

In der Literatur existieren eine Vielzahl von EKG-Signalmusterdatenbanken, z.B. (17–20). Die öffentlich verfügbaren Datenbanken basieren jedoch meist auf Datensammlungen für ausgewählte Befundgruppen. Die EKG-Meßdaten enthalten oft nicht alle 12 Ableitungen. Die Anzahl der verfügbaren EKG ist in vielen Fällen nicht ausreichend. Es konnte keine EKG-Datenbank gefunden werden, die den Anforderungen für das Mustererkennungsverfahren gerecht wird. Aus diesen Gründen war der Aufbau einer eigenen EKG-Signalmusterdatenbank notwendig.

Für den Aufbau der EKG-Signalmuster-Datenbank wurden ca. 12000 EKG von verschiedenen kardiologischen Schulen verwendet. Die messtechnische Signalqualität war z.T. sehr unterschiedlich. Für die Einbeziehung eines EKG in die Datenbank wurden folgende Mindestanforderungen gestellt:

- Bereitstellung des EKG als digitales 12-Kanal-EKG mit 10 s Messdauer (Ableitungen aVR, aVL, aVF, I, II, III, V1, V2, V3, V4, V5, V6) unter Berücksichtigung der Ableitungen Vx, Vy und Vz falls vorhanden
- Vorlage der Patientendaten Alter, Geschlecht, Größe und Gewicht
- Existenz eines EKG-Befundes.

Weitere verfügbare Informationen zu den technischen Parametern, Signalqualität, Signalbandbreite, Rauschen, Extraschläge, Offset, EKG-Rhythmus usw. wurden, falls vorhanden, erfasst.

Die EKG-Messdaten wurden ableitungsbezogen als 16-Bit-Binärfiles in einer Verzeichnisstruktur abgelegt. Jedem EKG wurde eine fortlaufende eindeutige EKG-Nummer als primärer Datenbankschlüssel zugewiesen. Ein zweiter sekundärer Datenbankschlüssel wurde für jeden Patienten (Patienten-Nummer) vergeben. Durch Dezimierungsfilterung wurde die Abtastrate der EKG-Ableitungen auf einheitlich 100 Hz reduziert. Technischer Hintergrund für diese Reduzierung ist die dadurch verminderte Rechenzeit bei der späteren EKG-Analyse.

Die mit der EKG-Signaldatenbank zusammenarbeitenden Programme gestatten es, jedes EKG aus dieser Verzeichnisstruktur aufzurufen, ableitungsweise oder als komplettes EKG darzustellen und auszugeben sowie für Signalverarbeitungsschritte einzeln oder im Stapel weiter zu verarbeiten.

Von ca. 10000 EKG wurde manuell ein repräsentativer EKG-Schlag über alle 12 bzw. 15 Ableitungen ausgewählt und mit weiteren, für die Korrelationsrechnung benötigten Zwischenergebnissen, wie z.B. der R-Zackenposition, in einer speziellen Rekordstruktur als File abgelegt. Entsprechend der Voraussetzungen für das Verfahren (6) wurden bei der visuellen Schlagauswahl ca. 1500 EKG ausgesondert und verworfen. Gründe hierfür waren:

- EKG mit Abweichungen des Schlag zu Schlag-Abstandes von größer 20%
- EKG mit verschiedenen, alternierend auftretenden Schlagtypen
- EKG mit sehr großen Basislinienschwankungen oder Drift einzelner Ableitungen
- EKG mit starkem Rauschen (P-Wellen im Rauschen nicht mehr erkennbar).

Tab. 2 Zuordnung Befundformulierung zu Wichtungsfaktoren (Konfidenzniveau)

Formulierung	Wichtungsfaktor (Konfidenzniveau)
– nicht auszuschliessen...	15%
– cannot rule out...	
– cannot be excluded...	
– möglicherweise...	35%
– consider...	
– suggest...	
– likely...	
– wahrscheinlich...	50%
– possible...	
– ...möglich	
– may be due...	
– probably due...	
– ablaufend(er)...	
– Verdacht auf...	
– Zustand nach...möglich	
– Sonst...	80%
– Bild...	
– Consistent with...	
– ...Diagnose	100%
– Zustand nach...	

Tab. 3 Zuordnung von Formulierungen zum Infarktstadium

Text Formulierung	Infarktstadium
– Acute	Stadium I
– Early	
– Stadium I – akut	
– Acute/subacute	Stadium I–II
– Ablaufend	
– Stadium I–II	
– Recent	Stadium II
– Subacut	
– Bereits abgelaufen	
– Stadium II	
– Subakut/chronisch	Stadium II–III
– Stadium II–III	
– Old	Stadium III
– Abgelaufen	
– Chronisch	
– Stadium III	
– Uncertain	Unbestimmt
– Unknown	
– Unbekannt	

Somit standen letztlich 8500 EKG für die weitere Auswertung zur Verfügung.

Die Patientendaten, die technischen Daten sowie die EKG-Befunde wurden in eine MS Access-Da-

tenbank übernommen. Die als fortlaufender Text vorliegenden EKG-Befunde konnten für eine computer-gestützte Auswertung in dieser Form nicht verwendet werden, da sie sowohl von verschiedenen Ärzten als auch in verschiedenen Sprachen erstellt wurden. Deshalb wurden die EKG-Befunde in Anlehnung an die „Acronyms for ECG interpretative Statements“ (9, 10) der American Heart Association (AHA) zerlegt. Dabei wurden der EKG-Befund in

- max. 5 diagnostische Aussagen (AHA1–AHA5)
- max. 3 Aussagen für die Signalform (FORM1–FORM3)
- max. 2 Aussagen für den EKG-Rhythmus (RHYTM1–RHYTM2)
- 1 Aussage zum Infarktstadium (ISTAD) bei Myokardinfarkten aufgespalten.

Für jede diagnostische Aussage des EKG-Befundes wurde außerdem eine Wahrscheinlichkeit für das Zutreffen dieser Aussage als prozentualer Wichtungsfaktor (Konfidenzniveau) angegeben. Dabei wurden in Anlehnung an bestehende EKG-Interpretationsprogramme bestimmte Schlüsselwörter des EKG-Befundes bestimmten Wichtungsfaktoren (Konfidenzniveau) zugeordnet (Tabelle 2).

Eine ähnliche Zuordnung erfolgte ebenfalls für das Merkmal Infarktstadium (Tabelle 3). Als Merkmale für die Infarktstadien I, I–II, II, II–III sowie III werden die Kriterien nach (11) zugrunde gelegt.

Ergebnisse der Auswertungen

Für die Analyse standen 8500 EKG zur Verfügung. Jedes dieser EKG wurde einzeln als unbekanntes Test-EKG gegen alle übrigen EKG als Datenbasis analysiert. Durch Datenbankabfragen wurden folgende EKG für diese statistischen Untersuchungen ausgeschlossen:

- Kinder-EKG
- Schrittmacher-EKG
- EKG mit Hinweisen auf möglicherweise vertauschte Elektroden

- EKG mit speziellen Elektroden-Anordnungen
- EKG ohne eindeutig zuordenbaren EKG-Befund
- EKG mit dem Befund „trace only requested (TOR)“.

Im Ergebnis der Rechnungen erhält man für jedes EKG eine prozentuale Verteilung der Hauptbefundgruppen bzw. der Unterbefundgruppen entsprechend der Beispiele Bild 2 bzw. 3.

Daraus wurden unter verschiedenen Analyseschwerpunkten die im Folgenden dargelegten Sensitivitäten und Spezifitäten berechnet.

Auswertung Normalpersonen

Eine gesunde Person wurde als „normal erkannt“ gewertet, wenn bei der Auswerttiefe von 50 EKG mindestens 35 EKG mit dem Befundstatement *NORM* gefunden wurde sowie Anzahl der EKG in den einzelnen pathologischen Befundgruppen (*MI*, *BBB*, *ISC*, *POSNL*) jeweils unter 10 lagen. Für diese Auswertung wurden 7570 EKG herangezogen (alle EKG mit Sinusrhythmus). Da die Befundgruppe Normal keine Untergruppen besitzt, müssen somit von den 50 in die Auswertung einbezogenen ähnlichsten EKG mindestens 35 den

Befund *NORM* (Normalbefund) besitzen.

Für die Hauptbefundgruppe Normal ergibt sich eine Sensitivität von 81% bei einer Spezifität von 95%. Bei vielen der nicht als *NORM* erkannten EKG wurde nicht die Zahl von 35 ähnlichen EKG mit dem Normalbefund gefunden. Eine mögliche Ursache dafür könnte in der großen biologischen Variabilität der Gruppe der Normalen und der möglicherweise noch nicht ausreichenden Zahl von Normalpersonen in der Datenbank bestehen. Senkt man die Zahl der für den Normalbefund erforderlichen ähnlichen EKG mit dem Befund *NORM* von 35 auf z. B. 30, steigt die Sensitivität auf 89% an, während die Spezifität auf nunmehr 92% zurückgeht.

Auswertung Myokardinfarkte

Ein Patient mit Myokardinfarkt wurde als erkannt gewertet, wenn bei einer Auswerttiefe von 50 EKG entweder

- mindestens 5 Befundhinweise *Myokardinfarkt* (*AMI*, *IMI*, *ILMI*, *IPMI*, *IPLMI*, *LMI*, *ASMI*, *PMI*, *IPMI*, *IPLMI*) gefunden wurden sowie der Befund Infarkt unter den 2 Befundhinweisen mit der größten Anzahl gefundener EKG erschien

Tab. 4 Sensitivität und Spezifität Befundgruppe Normalpersonen (*NORM*)

	NORM erkannt	NORM nicht erkannt		Sensitivität 0,81
NORM	2858	655	3513	
Nicht NORM	188	3869	4057	Spezifität 0,95
	3046	4524	7570	

Tab. 5 Sensitivität und Spezifität Myokardinfarkte (MI) ohne Befundwichtung

	MI erkannt	MI nicht erkannt		Sensitivität 0,66
MI	845	430	1275	
Nicht MI	859	5073	5932	Spezifität 0,86
	1704	5503	7207	

Tab. 6 Sensitivität und Spezifität Myokardinfarkte (MI) mit Berücksichtigung der Befundwichtung

	MI erkannt	MI nicht erkannt	
MI	407	157	564
Nicht MI	1297	5346	6643
	1704	5503	7207

Sensitivität
0,72Spezifität
0,80**Tab. 7** Sensitivität Myokardinfarkte (MI) mit Berücksichtigung der Befundwichtung sowie zusätzlich des Infarktstadiums

Infarkt-Stadium	MI erkannt	MI nicht erkannt	Anzahl	Sensitivität
Stadium I	36	12	48	0,75
Stadium I–II	76	8	84	0,90
Stadium II	70	18	88	0,80
Stadium II–III	10	7	17	0,59
Stadium III	171	102	273	0,63
Unbekannt	44	10	54	0,81

Tab. 8 Sensitivität Myokardinfarkte (MI) bei Berücksichtigung von 2 Infarktstadien

Infarktstadium	MI erkannt	MI nicht erkannt	Anzahl	Sensitivität
Stadium I–II	182	38	220	0,83
Stadium II–III	181	109	290	0,62

- oder mindestens 10 Befundhinweise *Myokardinfarkt* gefunden wurden sowie der Befund *Myokardinfarkt* unter den 5 Befundhinweisen mit der größten Anzahl gefundener EKG erschien.

trachtet und in die Untersuchung mit einbezogen. Insgesamt wurden für die Auswertung 7207 EKG verwendet (zusätzlich ausgeschlossen *AFIB*).

Damit ergibt sich für die Gruppe Infarkte eine Sensitivität von 66% bei einer Spezifität von 86%.

Auswertung ohne Befundwichtung

Alle 1275 EKG, bei denen im EKG-Befund der Begriff Infarkt als Befundstatement verwendet wurde, wurden unabhängig von der jeweiligen Befundwichtung (Konfidenzniveau) und unabhängig vom Infarktstadium für die Auswertung herangezogen. Das bedeutet, selbst wenn bei einem EKG nur mit 15%-iger Wahrscheinlichkeit ein alter Myokardinfarkt nicht ausgeschlossen werden konnte, wurde dieses EKG als Infarkt-EKG be-

Auswertung mit Befundwichtung

Es wurden für die Berechnungen nur die 564 EKG verwendet, bei denen im EKG-Befund der Begriff Infarkt als Befundstatement mit der höchsten Wichtung (Konfidenzniveau) jedoch unabhängig vom Infarktstadium vermerkt war. Aus Tabelle 5 ergibt sich damit eine Sensitivität von 72% bei einer Spezifität von 80%.

In einer weiteren Untersuchung wurde unter den gleichen Bedingungen wie bei der vorherigen Auswertung die Sensitivität in Abhängigkeit vom Infarktstadium untersucht. Die Ergebnisse sind in Tabelle 7 dargestellt.

Wie aus Tabelle 8 zu entnehmen ist, beträgt die Sensitivität in den Stadien I, I–II sowie II etwa 83%, während in den Stadien II–III und III noch 62% der Myokardinfarkte erkannt werden.

Auswertung Blöcke

Ein „Block“ wurde als erkannt gewertet, wenn bei einer Auswertetiefe von 50 EKG mindestens 10 Befundstatements *Block* (*CLBBB*, *CRBBB*, *LAfB*, *LPfB*, *IVCB*, *IVCD*, *IRBBB*, *ILBBB*) gefunden wurden sowie der Befund *Block* unter den drei Befundhinweisen mit der größten Anzahl gefundener EKG erschien. Für die Auswertung wurden 7521 EKG verwendet.

Auswertung ohne Befundwichtung

Insgesamt wurden 1581 EKG verwendet, bei denen im EKG-Befund der Begriff „Block“ als Befundstatement unabhängig von der jeweiligen Wichtung (Konfidenzniveau) vorkam. Unter diesen Bedingungen ergibt sich eine Sensitivität von 75% bei einer Spezifität von 92% für die Hauptbefundgruppe *Block*.

Auswertung mit Befundwichtung

Es wurden 1167 EKG, bei denen im EKG-Befund der Begriff „Block“ als Befundstatement mit der höchsten Wichtung (Konfidenzniveau) vorkam, verwendet. Hier ergibt sich für die Sensitivität ein Wert von 93% bei einer Spezifität von 91%.

Tabelle 11 zeigt die Sensitivität für die verschiedenen Untergruppen des EKG-Befundes *Block*. Betrachtet man nur die Blocktypen *CLBBB*,

Tab. 9 Sensitivität und Spezifität *Block* ohne Berücksichtigung der Befundwichtung

	Block erkannt	Block nicht erkannt		Sensitivität 0,75
Block	1191	390	1581	
Kein Block	448	5492	5940	Spezifität 0,92
	1639	5882	7521	

Tab. 10 Sensitivität und Spezifität *Block* mit Berücksichtigung der Befundwichtung

	Block erkannt	Block nicht erkannt		Sensitivität 0,93
Block	1084	83	1167	
Nicht Block	555	5799	6354	Spezifität 0,91
	1639	5882	7521	

Tab. 11 Sensitivitäten der Untergruppen des Befundes *Block* mit Berücksichtigung der Befundwichtung

	Block erkannt	Block nicht erkannt	Sensitivität
CRBBB	296	7	0,98
CLBBB	151	3	0,98
LAFB	511	13	0,98
IVCB	32	10	0,76
ILBBB	4	4	0,50
IRBBB	73	39	0,65
LPFB	3	3	0,50
IVCD	14	4	0,78
	1084	83	1167

Tab. 12 Sensitivität und Spezifität *Ischämie* ohne Berücksichtigung der Befundwichtung

	ISC erkannt	ISC nicht erkannt		Sensitivität 0,80
ISC	1317	327	1644	
Nicht ISC	895	4983	5878	Spezifität 0,85
	2212	5310	7522	

Tab. 13 Sensitivität und Spezifität *Ischämie* mit Berücksichtigung der Befundwichtung

	ISC erkannt	ISC nicht erkannt		Sensitivität 0,86
ISC	1170	186	1356	
Nicht ISC	1042	5124	6166	Spezifität 0,83
	2212	5310	7522	

CRBBB, *LAFB*, *IVCB*, erhält man für diese Befundgruppe eine Sensitivität von 97%. Relativ schlechte Sensitivitäten erhält man dagegen für die Befundgruppen *IRBBB*, *ILBBB* sowie *LPFB*.

Auswertung Ischämien

Ein Befund „Ischämie“ wurde als erkannt gewertet, wenn bei einer Auswertetiefe von 50 EKG mindestens 10 Befundaussagen *Ischämie* gefunden wurden sowie der Befund „Ischämie“ unter den ersten drei Befundhinweisen mit der größten Anzahl gefundener EKG erschien. Für die Auswertung wurden 7522 EKG verwendet.

Auswertung ohne Befundwichtung

Alle 1644 EKG, bei denen im EKG-Befund der Begriff „Ischämie“ als Befundstatement unabhängig von der Wichtung (Konfidenzniveau) verwendet wurde, wurden für die Auswertung herangezogen.

Im Ergebnis der Auswertung erhält man für die Befundgruppe Ischämie eine Sensitivität von 80% bei einer Spezifität von 85%.

Auswertung mit Befundwichtung

Es wurden für die Berechnungen 1356 EKG verwendet, bei denen im EKG-Befund der Begriff „Ischämie“ als Befundaussage mit der höchsten Wichtung (Konfidenzniveau) auftrat.

Hierbei ergibt sich für den Befund „Ischämie“ eine Sensitivität von 86% bei einer Spezifität von 83%.

Zusammenfassung der Auswertungen

Bild 4 und 5 zeigen zusammenfassend die erreichten Ergebnisse für Sensitivität und Spezifität der Hauptdiagnosegruppen bei Berück-

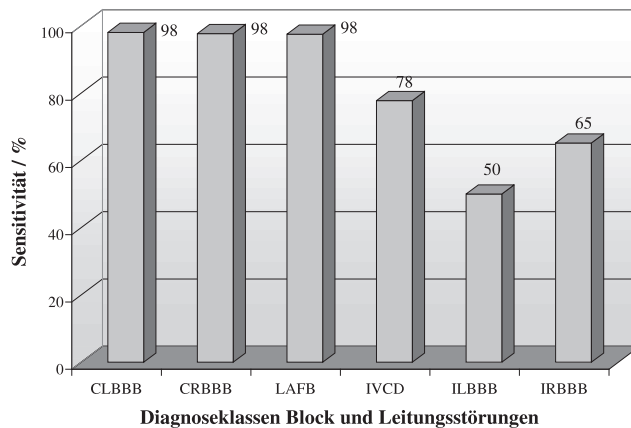


Abb. 5 Sensitivität der Erkennung der verschiedenen Blocktypen

sichtigung der Befundwichtigungen (Konfidenzniveau) sowie der Unterbefundgruppen Block.

Insgesamt ergibt die Berücksichtigung des Konfidenzniveaus eine verbesserte Sensitivität bei etwa gleichbleibender bzw. nur leicht zurückgehender Spezifität.

Diskussion

Verfahrensbedingt kommt bei dem Mustererkennungsverfahren der Qualität der Befundhinweise der Datenbank-EKG eine entscheidende Bedeutung zu. Die im Ergebnis der Anwendung des Verfahrens erzielten Werte für die Sensitivität und die Spezifität hängen sehr stark von der Richtigkeit und Vollständigkeit der EKG-Befunde sowie deren Zerlegung in die entsprechend richtig gewichteten Befundaussagen ab. Weiterhin ist die Widerspiegelung des biologischen Spektrums kardiologischer Erkrankungen und der gesunden Vergleichspersonen eine Voraussetzung dafür, dass vergleichbare EKG in der Datenbank gefunden werden können. Da zu den meisten EKG-Befunden keine weiteren validierenden Untersuchungsergebnisse vorlagen, kann davon ausgegangen werden, dass die EKG-Befunde eine nicht unwesentliche, vom jeweils

befundenden Kardiologen abhängige subjektive Komponente enthalten.

Durch die große Anzahl der EKG, die unterschiedliche Herkunft sowie die Befundung durch verschiedene Kardiologen wird diese subjektive Komponente zwar relativiert, ist jedoch als entscheidende Voraussetzung für die weitere Verbesserung der Interpretationsergebnisse anzusehen. Mögliche Fehlerquellen bei den in der EKG-Datenbank enthaltenen Befunde sind:

- Einer oder mehrere der in AHA1–5 aufgeführten Befunde sind falsch (siehe Beispiele Tabelle 14)
- Die Wichtungsfaktoren der Befunde (Konfidenzniveau) sind falsch (wurde dem Hauptbefund die höchste Wichtung zuerkannt?).

Die Klärung dieser Probleme kann nur in enger Zusammenarbeit mit Kardiologen erfolgen.

Bedingt durch den Entwicklungsstand des Algorithmus sind die nachfolgend aufgeführten, möglichen Fehlereinflüsse auf die vorgestellten Ergebnisse zwar bekannt, können jedoch noch im Einzelfall zu Interpretationsfehlern führen:

1. Der Signalmustervergleich des Test-EKG mit den EKG der Referenz Datenbank führt nicht zu Korrelationsergebnissen im Ziel- oder Auswertebereich. Ursachen dafür können sein:

- Keine ähnlichen EKG in der Referenzdatenbank vorhanden.
 - Starke Signalmusterverzerrungen durch Drift in einer oder mehreren Ableitungen des Test-EKG.
 - Signalmusterabhängige Fehler bei der richtigen Bestimmung der verfahrensbedingt erforderlichen R-Zacke-Position über alle Ableitungen des Test-EKG sowie der Datenbank-EKG.
2. Subjektive Auswahl eines für das EKG typischen Schläges.
 3. Der gewählte typische EKG-Schlag wurde von starken Störungen in Signalabschnitten innerhalb bzw. außerhalb des Auswahlfensters überlagert (z.B. Rauschen).
 4. Aus Gründen der Rechenzeit erfolgt die Korrelationsanalyse gegenwärtig mit EKG-Signalen der Abtastfrequenz 100 Hz. Kleine Amplituden-Zeit-Veränderungen im Signalmuster des EKG-Signals mit großer diagnostischer Relevanz werden damit möglicherweise unzureichend berücksichtigt.

Unabhängig von den bereits erkannten Fehlerquellen wurden *alle* zur Verfügung stehenden EKG zur Auswertung verwendet. Die Auswertung erfolgte somit auch in den Fällen, in denen aufgrund fehlender Korrelationen im Auswertebereich entsprechend den Voraussetzungen für das Verfahren keine sichere Auswertung zulässig ist.

Die unter diesen Voraussetzungen erhaltenen Ergebnisse für die Sensitivität und die Spezifität der untersuchten Befundgruppen stellen also keineswegs die bestmöglichen Ergebnisse dar, die sich mit Hilfe der Methode der Mustererkennung erreichen lassen. Trotzdem wurden mit der vorliegenden Datenbank, auch unter Berücksichtigung der diskutierten Fehlerquellen, Ergebnisse erreicht, die z.T. über den Empfehlungen für rechnergestützte EKG Interpretationsprogramme liegen (12, 13).

Tab. 14 Beispiele für die Überprüfung der EKG-Befunde der Datenbank

EKG-Nr.	Befundtext der EKG-Datenbank	Analyseergebnis nach dem neuen Verfahren
1174	Sinus arrhythmia. normal ecg.	Incomplete bundle branch block (IBBB) + left anterior fascicular block (LAFB)
1710	Sinus rhythm. Prolonged pr interval. Otherwise normal ecg.	Left anterior fascicular block (LAFB)
7384	Sinus rhythm. normal ecg.	Inferiorer myocardial infarktion (IMI), Stadium III

Tab. 15 Vergleich der Sensitivität und der Spezifität des Mustererkennungsverfahrens mit den „Performance requirements for recording and analyzing electrocardiographs (12), 13“

Diagnose-Gruppe	Mustererkennungsverfahren		Performance requirements (12, 13)	
	Sensitivität	Spezifität	Sensitivität	Spezifität
Norm	89%	92%	85%	70%
INF	72%	80%	65%	85%
CLBBB	98%	91% *	80%	85%
CRBBB	98%	91% *	80%	85%

* Spezifität für die gesamte Gruppe Blöcke (CLBBB, CRBBB, LAFB, LPFB, IVCB, IVCD, IRBBB, ILBBB)

Zusammenfassung

Zielstellung der Untersuchungen war der Nachweis der Funktion des Mustererkennungsverfahrens an einer großen, weitgehend das gesamte Spektrum kardiologischer Erkrankungen umfassende EKG-Datenbank. Im Vergleich zum ersten Funktionsnachweis des Verfahrens (6, 14) wurde das Mustererkennungsverfahren hinsichtlich seiner Rechengeschwindigkeit weiter optimiert und eine EKG-Datenbank mit insgesamt 12000 EKG aufgebaut. Es konnte gezeigt werden, dass ein Mustervergleich eines unbekannten EKG mit einem Satz von 8500 Referenz-EKG auf einen normalen PC (400 MHz Pentium II) in 20–30 s durchführbar ist.

Im Unterschied zu den bisherigen Untersuchungen (6–8, 14) mit den Diagnosegruppen Normale und Infarkte, konnte nachgewiesen werden, dass auch mit einer großen, das gesamte Spektrum kardiologischer Befunde umfassenden Datenbank eine gute Trennung der untersuchten Hauptbefundgruppen möglich ist. Bei der Analyse von realen EKG werden in den Befundgruppen Norm, Block, Infarkt und Ischämie Sensitivitäten von 72–98% (Mittelwert 83%) bei Spezifitäten von 80–95% (Mittelwert 87%) erreicht. Dies trifft analog auch für die Trennung der untersuchten Unterbefunde z.B. bei Blöcken und Infarkten zu. Aufgrund der ungünstiger Ausgangsbedingungen (Validierung der EKG-Befunde, methodische Fehler) lassen sich diese Werte weiter verbessern.

Ein wesentlicher Vorteil des Mustererkennungsverfahrens gegenüber bisherigen Verfahren besteht in der Berücksichtigung von EKG mit Mehrfachbefunden. Bedingt durch die spezifischen Signalmuster, vergleichbar mit einem „Fingerabdruck“, lassen sich auch bei schwierigen Fällen ähnliche EKG finden und damit statistisch gesicherte Hinweise auf den Befund ableiten.

Durch die Trennung des Analyseverfahrens in einen unveränderlichen algorithmischen Teil und in eine austausch- und erweiterbare EKG-Signaldatenbank lassen sich neue Versionen von EKG-Analyseprogrammen durch einfachen Wechsel der zugrunde liegenden Datenbasis schaffen. Mögliche Perspektiven der weiteren Entwicklung könnten z.B. noch umfangreichere, gut validierte EKG-Datenbanken sein, die auch im WWW verfügbar sind. Auch der Aufbau spezialisierter EKG-Datenbanken, z.B. für die Analyse von Kinder-EKG oder die Berücksichtigung spezieller Patientengruppen, ist denkbar.

Bereits funktionsfähig ist eine netzwerkbasierte Lösung, die es ermöglicht, von einem EKG-Aufnahmesystem (z.B. Laptop mit PC-EKG-Gerät) über das Intra- bzw. das Internet EKGs auf einem entfernten Auswertesystem analysieren zu lassen. Bei dieser Lösung auf der Grundlage der Client-Server-Technologie sind die EKG-Aufnahme und die Auswertung und die Ergebnisdarstellung räumlich getrennt. Dies ermöglicht den Aufbau leistungsfähiger zentraler EKG-Datenbank- und Analysesysteme und kostengünstiger EKG-Aufnahmesysteme.

Literatur

- Macfarlane PW, Devine B, Latif S, Maclaughlin SC, Shoat DB, Watts MP (1990) Methodology of ECG interpretation in the Glasgow Program, *Methods of Information in Medicine* 29:354–361
- Gebrauchsanweisung EKG-Vermessungsprogramm und Interpretationsprogramm. Ausgabe 1997, Schiller AG, Baar, Schweiz
- Siemens-Kriterien für die EKG-Auswertung (1988) Siemens-Elementa AB. Medicinsk Teknik, Solna, Sweden
- Willems JL (1977) Introduction to multivariate and conventional computer ECG analysis: pros and cons. *Trends in Computer-Processed Electrocardiograms*, Amsterdam, pp 213–220
- Bortolan G, Degani R, Willems JL (1991) ECG classification with neural networks and cluster analysis. *Proceedings Computers in Cardiology* (cat. No. 91CH3116-1). Held: Venice, Italy, 23–26 Sept. 1991. (USA: IEEE Comput Soc Press, pp 177–80)
- Bousseljot R, Kreiseler D (1998) Auswertung von EKG mit Hilfe der Mustererkennung Herzschrittmachertherapie & Elektrophysiologie. *Steinkopff* 9:270–278
- Kreiseler D, Bousseljot R (1998) EKG-Signalanalyse durch Mustervergleich und EKG-Datenbanken. *Biomedizinische Technik*, Band 43, Ergänzungsband 3
- Bousseljot R, Kreiseler D (1998) ECG Signal Analysis by Pattern Comparison *Computers in Cardiology*, Vol 25, S 349–352
- Standardization of Terminology and Interpretation (1978) *American Journal of Cardiology* 41:130–145
- Definitions and Classification of Cardiac Arrhythmias (1978) *American Heart Journal* 95:796–806
- Netter FH (1976) *Farbatlanten der Medizin, The Ciba Collection of Medical Illustrations*. Band 1, Herz, Georg Thieme Verlag Stuttgart
- Willems J, Arnaud P, Van Bommel J, Degani R, Macfarlane PW, Zywiets C (1990) Common standards for quantitative electrocardiography – goals and main results, *Methods of Information in Medicine* 29:263–271
- Willems J, Abreu-Lima C, Arnaut P, Van Bommel J, Brohet C, Degani R, Denis B, Gehring J, Graham I, Van Herpen G, Machado H, Macfarlane P, Michaelis J, Mouloupolous SD, Rubel P, Zywiets C (1991) The Diagnostic Performance of Computer Programs for the Interpretation of Electrocardiograms, *New England Journal of Medicine* 325:1767–1833
- Bousseljot R, Kreiseler D (1998) Ergebnisse bei der EKG-Analyse durch Signalmustervergleich. *Biomedizinische Technik*, Band 43, S 234–235
- Bousseljot R, Kreiseler D (1996) Auswertesystem zur Gewinnung diagnostischer Informationen aus Signalen und Daten medizinischer Sensorsysteme DE-OS 196 38 738 A 1, A 61B5/00, AT.: 10.09. 1996, OT 12.03. 1998
- Potschinkov A (1996) Universelle Filter zur Abtastfrequenzänderung. Ingenieurbüro Dr.-Ing. Alexander Potschinkov, Clausewitzstr 4, 10629 Berlin
- MIT-BIH Arrhythmia Database CD-ROM, Harvard-MIT Division of Health Sciences and Technology, Beth Israel Hospital, Boston
- AHA Database for Evaluation of Ventricular Arrhythmia Detectors, American Heart Association (AHA), National Heart, Lung and Blood Institute (NHLBI)
- CSE-Diagnostic-Database INSERM Unit 121, Hospital Cardiologique, 59 Boulevard Pinel, BP Lyon-Montchat, 69394 Lyon Cedex 3, France, z.H. Ph.D. Paul Rubel
- European ST-T-Database, National Research Council (CNR) Institute of Clinical Physiology