

CS5540: Computational Techniques for Analyzing Clinical Data

Prof. Ramin Zabih (CS)

Prof. Ashish Raj (Radiology)

Today only:

Gary Dorfman, MD (Radiology)

What is CS5540 about?

- The practice of medicine is being transformed by computation
- Information about patients is (rapidly) becoming digital and (slowly) becoming available for computers to analyze
- This opens huge opportunities for both CS and medicine
- CS5540 will cover CS techniques and their application to clinical data
 - Interest in medicine required; ignorance OK!



What might computers do?

- Analyze a patient's vital signs and automatically delivers treatment
- Automatically determine that a drug for inflammation increases heart attack risk
- Determine that a patient is suffering from epilepsy
- ETC

Autonomous treatment

- Sci-Fi fantasy: computer diagnoses and treats the patient all by itself
 - Especially when the treatment is a major intervention
- Suppose the treatment is lifesaving if the patient has one condition, but deadly if they have a different one
 - The computer diagnosis had better be right!
 - Could such a device exist this century?
- It does, and you've seen it

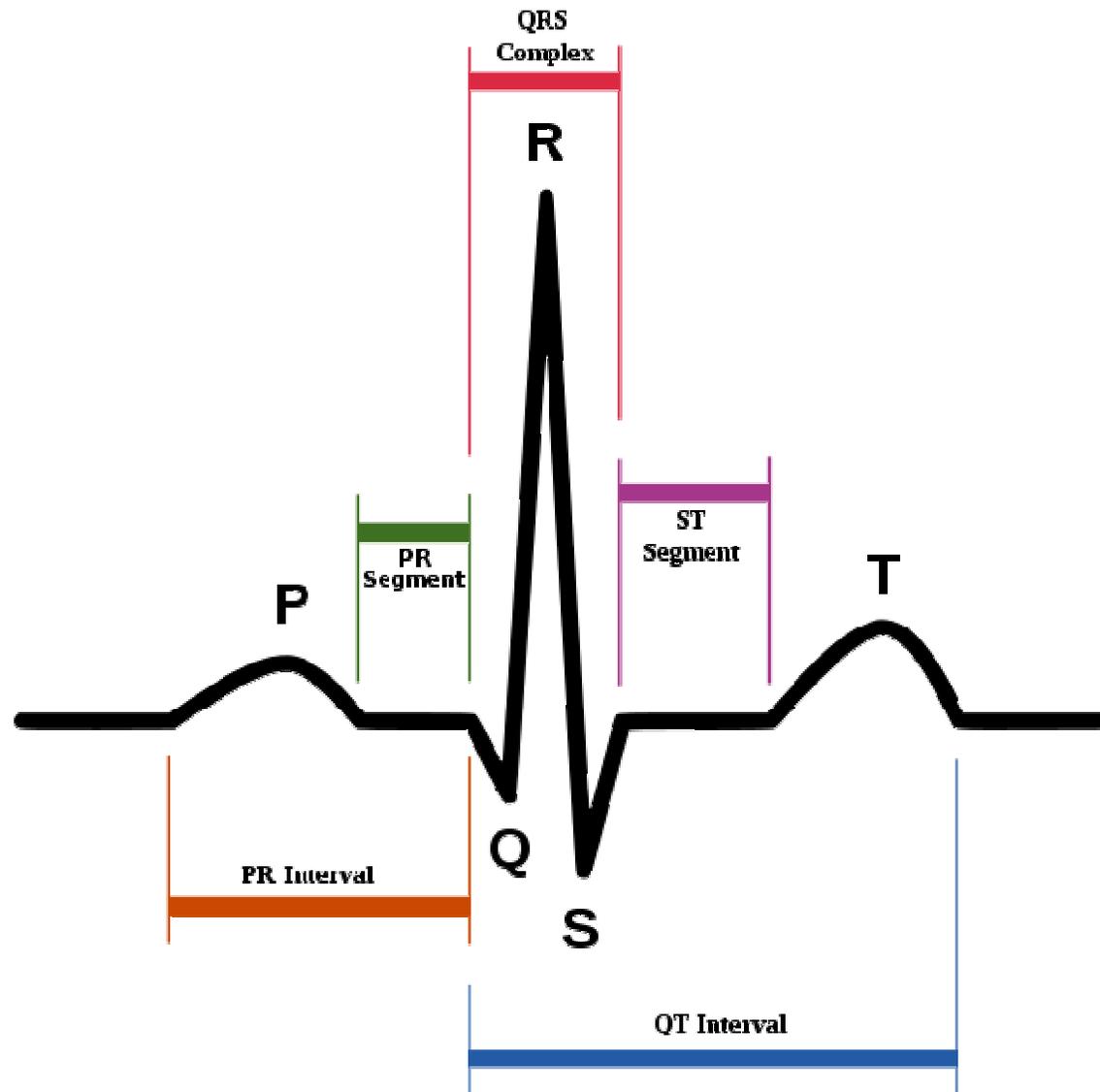




Automatic External Defibrillator



ECG of an idealized heartbeat



AED diagnosis

- The AED must quickly determine if the patient has a shockable rhythm
 - Not all arrhythmias are shockable
- Shocking a healthy patient can induce cardiac arrest
 - To publicly distribute this device, it needs to be resistant to stupid pranks (hazing, etc.)
 - It's amazing this is allowed (Gary Dorfman will describe some of the approval issues)



Determining drug side effects

- Drug approval is a laborious and expensive process (clinical trials)
- Can't cover all possible contingencies
 - Combinations of medications, underlying conditions, surgical interventions, etc.
- Sometimes nasty surprises emerge only after a drug is in widespread use
 - Famous examples?
 - Can be both deadly and extremely expensive



VIOXX saga

- Anti-inflammatory without GI upset
 - Based on blocking the COX2 enzyme
- After widespread release, it became clear that VIOXX increased cardiac risk
 - Careful retrospective analysis of the data
 - Many believe that the manufacturer (Merck) became aware of this during clinical trials
- Can we detect this automatically from patient records?
 - There is some evidence this is doable



Epilepsy diagnosis by imaging

- Many disorders give rise to seizures
 - With different treatments and prognoses
- Quantitative measures from images
 - Changes in thickness from MR
- Ideally this should be like “heart rate” or “red blood cell count”
 - A number, where normal is a certain range
 - A large population of patients known to have epilepsy lie outside this range
- “Imaging as a biomarker”



Common themes

- Computer-aided diagnosis
 - Usually not autonomous
- Treatment monitoring
 - Surprisingly important
- Wide variety of data, and lots of it
- Creation and analysis of biomarkers



Computational Techniques Applicable to Medical Data: One Clinician's View

Gary S. Dorfman, M.D.

**Professor and Vice-Chairman for Research
Department of Radiology
Weill Cornell Medical College**

Financial Disclosures

Dr. Dorfman serves on the Medical Advisory Board of Vascular Solutions, Inc.

Dr. Dorfman holds patents assigned to a company of which he is the sole owner and officer.

In the opinion of the presenter, neither of these disclosures is relevant to the material included in this presentation.



Medical Data Typology

- **Qualitative**
 - **Text – History, PE, Assessments, Plans, Results, Reports, etc.**
 - **Photographs of Lesions, Pathology Slides, etc.**
- **Semi-Quantitative**
 - ***in vivo* Imaging (Radiology)**
 - **Linear Data (ECG, EEG, EMG, etc.)**
- **Quantitative**
 - **Numerical Lab Results**
 - **Extracted Numerical Values from Qualitative and Semi-Quantitative Data (e.g., Mark-ups, Annotations)**
- **Raw vs. Processed vs. Post-Processed**



Medical Data Opportunities

- **Volume (Workload, Needle(s) in a Haystack)**
- **Storage (What, Where, How Long, Format?)**
- **Organization**
- **Presentation**
- **Extraction of “New” or “Hidden” Data**
- **Communication**
- **Query and Analysis**
- **Security**
- **Sharing**
- **Tools & Techniques, Environment & Ergonomics**



Administrivia

- Some programming experience is needed
 - Matlab and C are preferable, Java is OK
 - Some exposure to algorithms & statistics are helpful but not mandatory
 - Provide a deeper understanding of some topics
- Lectures by Prof. Zabih or Prof. Raj
 - Some guest appearances by physicians, typically on Wednesday (CS7594)
 - A few guest lectures by physicians



Main topics

- 1D data: ECG, EEG, EMG, etc
 - Analyzing repetitive waveforms
- 2D data: CT, MRI, PET, etc
 - Analyzing images
 - Creating high-quality images
- Other data: EMR, etc
 - Finding clinically important correlations



Algorithms

- Dynamic programming
- Graph algorithms, esp. min cut
- Fitting via least squares & its variants
- Gradient descent, conjugate gradient, PCG
- k-NN, SVM classification



Workload

- 2 or 3 projects
 - “Implement this algorithm we discussed in class, try it on these data sets, make some modifications”
 - Can be done in pairs, in any reasonable language (no Fortran/Postscript!)
- Small in-class quizzes every 2 weeks
 - Intended to be easy, short-answer
- Free form final project (express yourself!)
- No prelims or final exam



Analyzing 1D signals

- Many clinical measurements produce time-varying data that is roughly periodic
 - Classic example: electrocardiogram (ECG)
 - Lots of other examples from multiple fields
- How to understand an ECG?
- One way to proceed is *model-based*
 - Look at the electrophysiology of the heart
 - Understand the sources of measurement errors
- We're not going this route
 - Too specific, and effectiveness is unclear

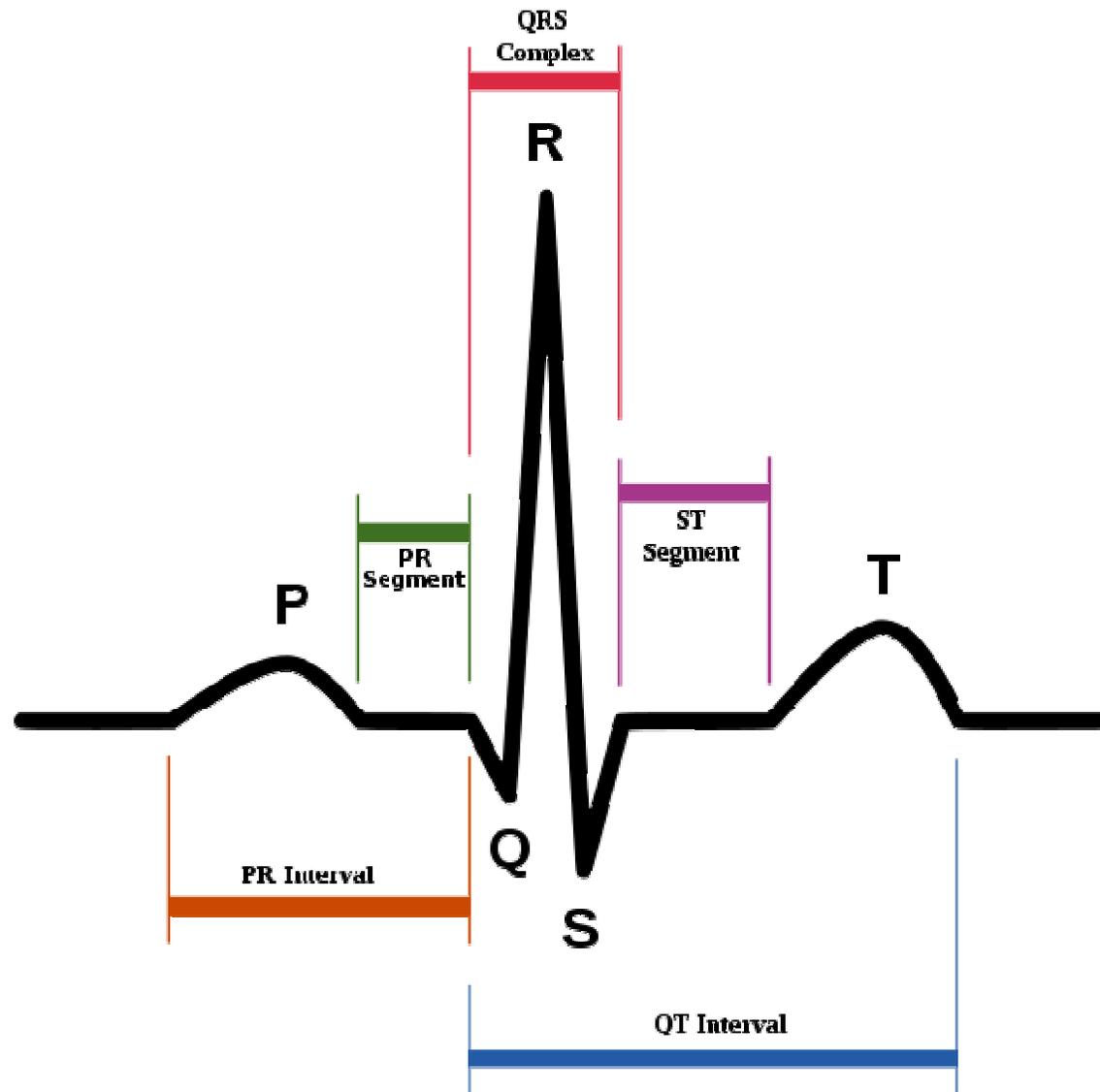


Tools for periodic 1D data

- Instead we will look at some general methods for roughly periodic 1D data
 - “Roughly” is important
- As a simple case, how does a patient’s ECG change over time?
 - For instance, as he runs on a treadmill, or holds his breath, or gets a medication
- What does a normal ECG look like?
- “Simple” question: what is the heart rate?



ECG of an idealized heartbeat



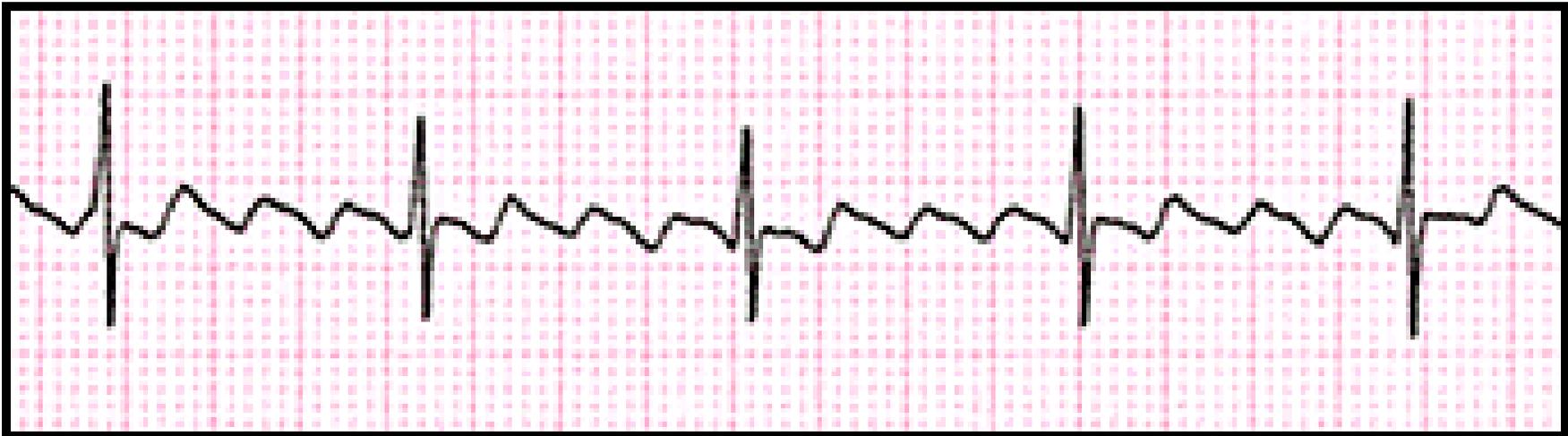
Sometimes, anything works



www.uptodate.com



Slightly harder case



www.uptodate.com



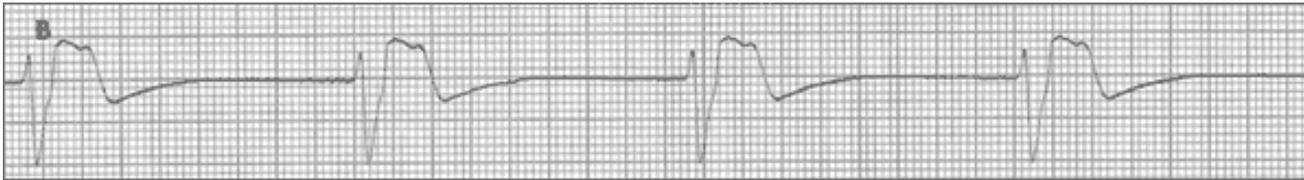
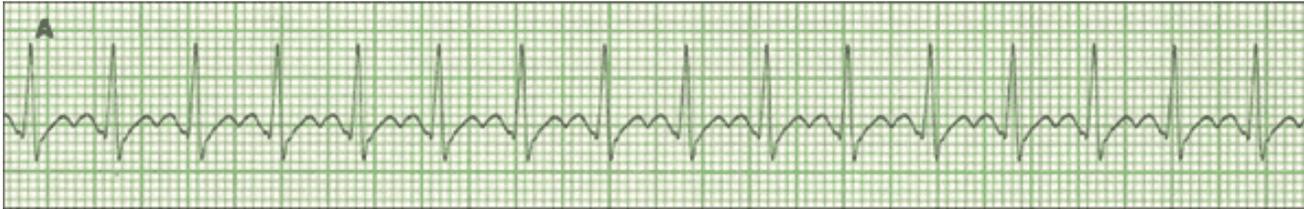
Life isn't always easy



The Alan E. Lindsay ECG Learning Center ; <http://medstat.med.utah.edu/kw/ecg/>



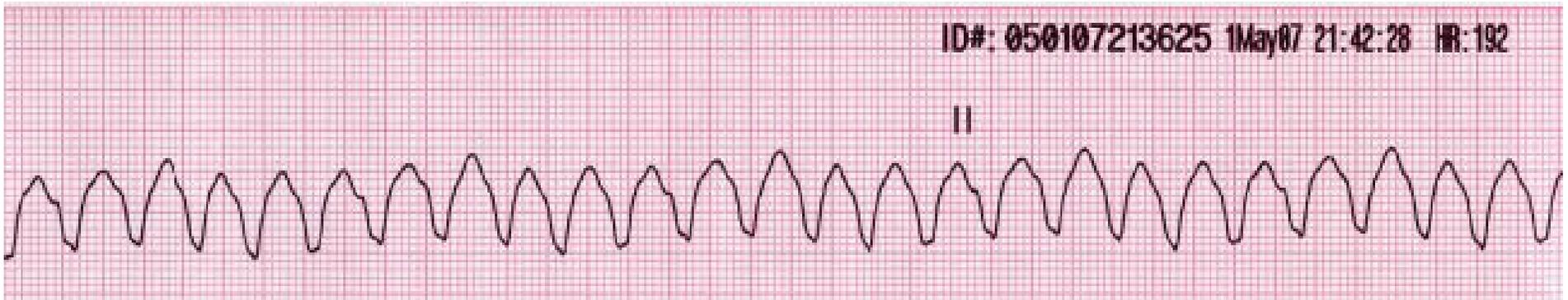
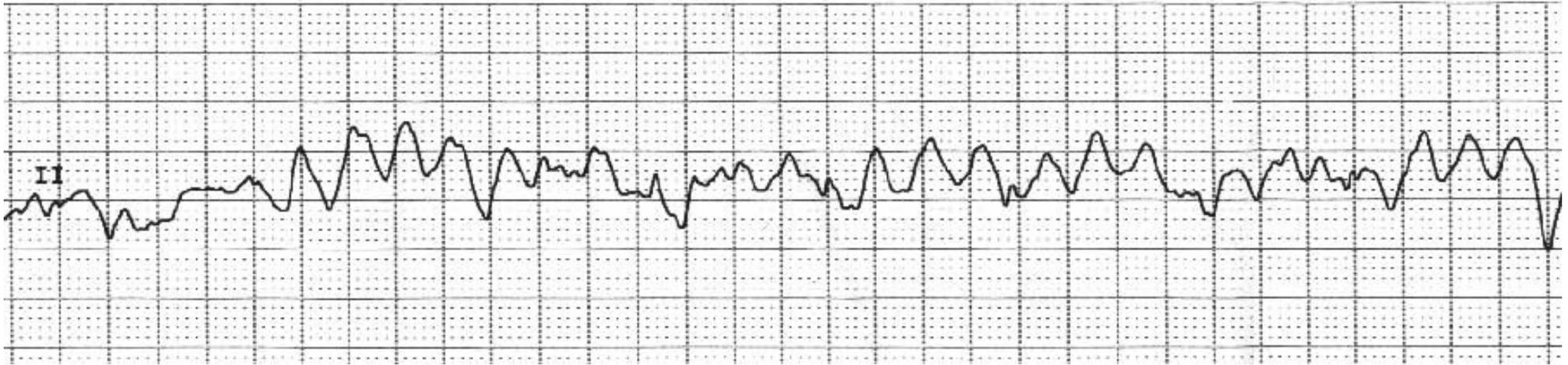
Non shockable arrhythmias



Brady, Harrigan and Chan, A Tale of Two “Pulseless Electrical Activity” Cardiac Arrest Rhythms, *Consultant* 47(1)



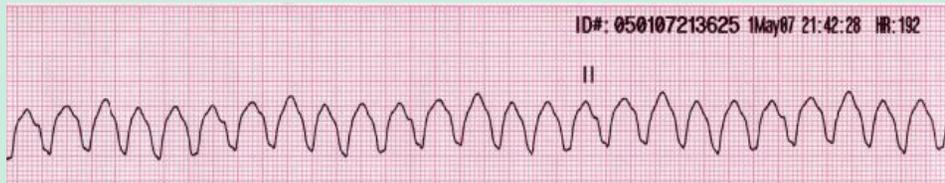
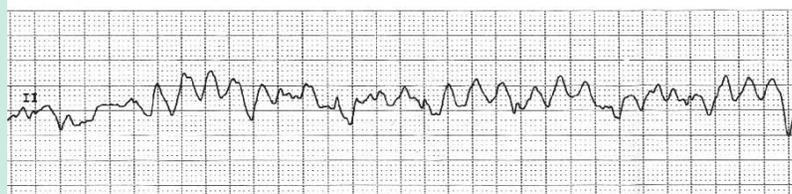
Shockable arrhythmias



Source: Wikipedia

Classification for an AED

Shockable



Non shockable



AED accuracy

- AHA goal: 90% sensitivity for V-Fib, 99% specificity for NSR
 - 10% false negatives for V-Fib
 - 1% false positives for NSR
 - Kerber et al., *Circulation* 95(6):1677-82, 1997



How to look for ECG changes?

- Three issues:
 - Divide the data into cycles
 - Compensate for speed differences
 - Analyze each cycle
- Important ideas
 - Segmentation
 - Alignment
 - Analysis

