

# Communication of Uncertainty in Clinical Genetics Patient Health Communication Systems\*

**Nancy Green, Tami Britt, and Karen Jirak**

University of North Carolina at Greensboro  
nlgreen @ uncg.edu

\*This material is based upon work supported by the National Science Foundation  
under CAREER Award No. 0132821

*AAAI 2004 Fall Symposium on Dialogue Systems for Health Communication*

## Long-term Goals

- **Motivation:** presentations of vital scientific or medical information may be difficult for lay audience to grasp, due to lack of familiarity with
  - diagnosis and other forms of causal reasoning, complex arguments
  - use of statistical evidence, basic probability
- **Proposal**
  - use AI to produce lay-oriented presentation of information (text, dialogue, graphics, interactivity)
  - use HCI methods to ensure effectiveness
  - embody research in a demonstration system (**GenIE**) for clinical genetics

## Why Clinical Genetics?

As knowledge about human genome increases, so have genetic testing options:

- newborn screening, other testable genetic disorders
- predisposition for cancer (BRCA1, etc.)
- variation affecting an individual's response to medications

Requires communication with lay audience on issues in biomedical domain of vital importance

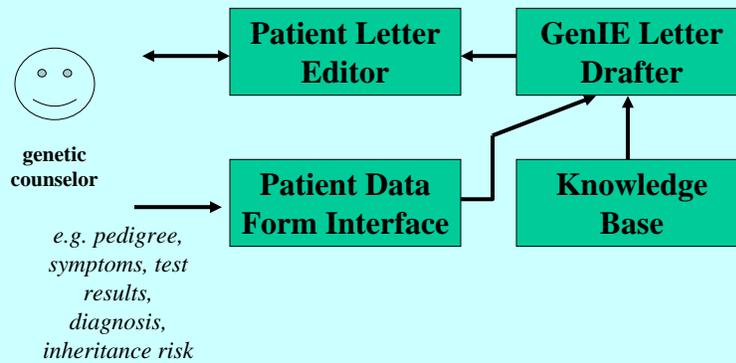
Currently most of this kind of communication is done by genetic counselors

## Today's Genetic Counselor

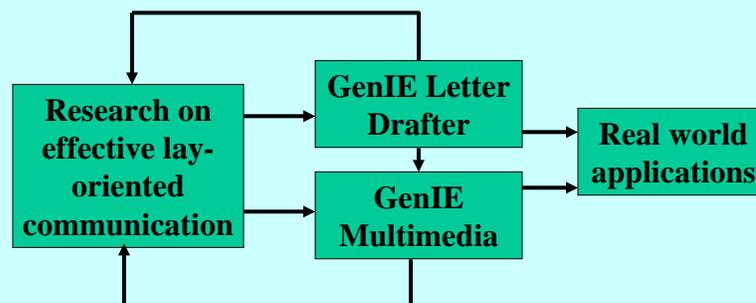
- Meets with clients of genetics clinic
  - Informational role (non-directive), e.g.
    - Explanation of diagnosis of genetic condition
    - Explanation of inheritance risks in family
    - Discussion of treatment/management of condition
  - Educational role
  - Supportive role (due to loss, perceived stigma, etc.)
- Writes a *patient letter* to client (1-2 pages)
  - Documentation for medical record
  - Client uses (e.g. communicate with family members)

## GenIE System (v.1)

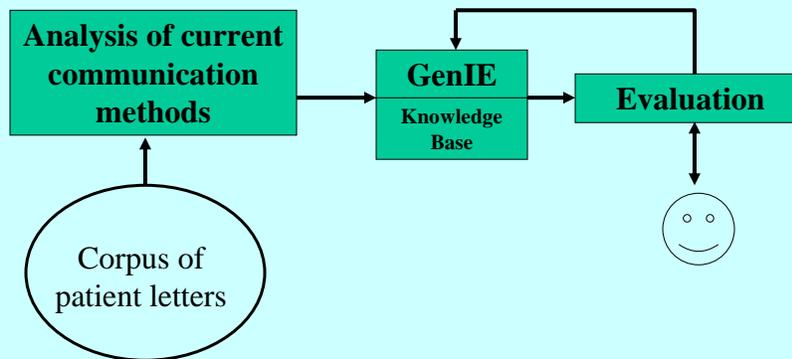
Goal to draft 1st draft of patient letter for genetic counselors



## R&D Strategy



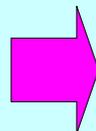
## Research on effective lay-oriented communication in clinical genetics



## Communication of uncertainty

A major problem area for effective communication with lay audiences in clinical genetics ...

- Challenges (see paper)
  - pervasive in genetics
  - issues in communication of probability



- **Uses of uncertainty in corpus**
- **Some implications for design of GenIE**

## Analysis of Uncertainty in Corpus

Began by manually tagging concepts in letters using coding scheme

[Green, N. A Bayesian Network Coding Scheme for Annotation of Biomedical Information Presented to Genetic Counseling Clients. To appear in *Journal of Biomedical Informatics*.]

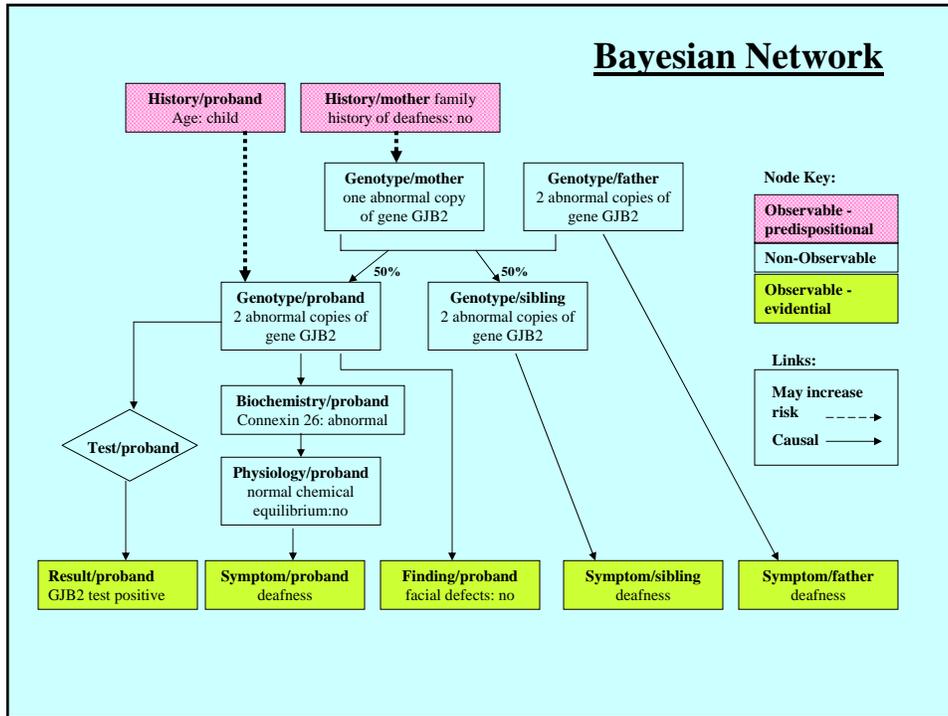
*Example: (7) Approximately 80% of individuals affected with NF have mild to moderate symptoms ...*

<probability-7 Approximately 80% >  
of individuals affected with  
<genotype-7 NF> have  
<symptom-7 mild to moderate symptoms>

## Analysis of Uncertainty in Corpus (2)

Role of coding scheme

- classify content at high level of abstraction: *genotype*, *symptom*, *test result*, etc.
  - represent commonalities among genetic processes
  - enables generalizations about communication strategies independently of particular genes, proteins, etc.
  - intercoder reliability evaluation: very good (Green, to appear, JBI.)
- represent causal & probabilistic relations communicated (implicit and explicit) in letter
- result of encoding is partially specified **Bayesian Network (BN)** for each letter



## Encoding of Probability Statements

In addition to Bayesian Network, can represent probability statements made in letter, e.g.:

(7) *Approximately 80% of individuals affected with NF have mild to moderate symptoms ...*

<probability-7 Approximately 80% >  
of individuals affected with  
<genotype-7 NF> have  
<symptom-7 mild to moderate symptoms>

$$P(\text{symptom-7} \mid \text{genotype-7}) = \text{probability-7}$$

## Features of Probability Statements

- **Retrospective**  
P(observable | diagnosis ) = probability  
or **Predictive**  
P(diagnosis | observable ) = probability
- **Progressive** (text in causal order)  
or **Regressive** (text in reverse causal order)
- **Quantitative value** (exact, qualified, or range)  
e.g. *80%, about 80%, 50% to 80%*  
or **Qualitative value:**  
e.g. *may, usually, possible, many, expected*

## Distribution in representative samples

In sample of 3 letters (102 sentences total):

- Average ratio of probability statements to sentences: 73 to 102
- Average ratio of retrospective to predictive: 30 to 42
- Average ratio of regressive to progressive: 24 to 28

In one other letter (46 sentences total):

- Quantitative (exact numeric) values: 5
- Qualified numeric and range values: 5
- Qualitative values: 29

## Uses of Uncertainty as Probability Value

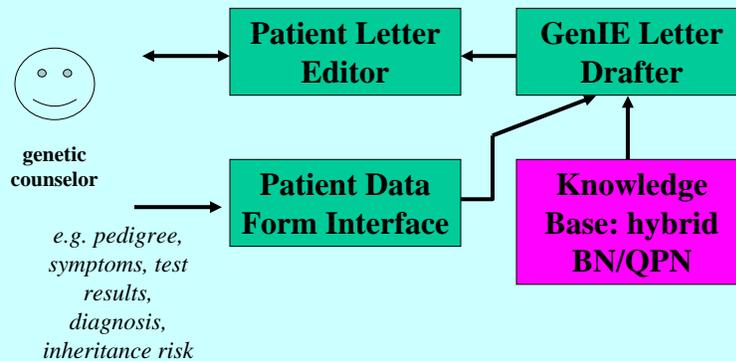
- **Evidence for hypothesis from population**  
*Individuals with VCF often have heart defects ...*
- **Evidence for hypothesis from patient data**  
*.. results of the collagen studies is consistent with your having an altered gene for ...*
- **Prediction of effect based on model (Mendelian theory)**  
*.. have a one in four (25%) chance that a child ..*
- **Comparison of likelihood**  
*The most common change in this gene is ..*
- **Base rate from population data**  
*.. has about a 2 to 3% (2 to 3 out of 100) chance of being a carrier ..*

## Rhetorical Uses of Uncertainty

- Restatement: 25% will occur => 75% will not occur
- Avoiding commitment
- Avoiding numeric overload (qualitative uses)
- Mitigating negative impact

## GenIE System

*Investigating use of qualitative probabilistic networks (QPN) in design of system's KB*



## Hybrid QPN/BN in GenIE

### Qualitative Probabilistic Network (QPN)

- abstraction of BN where numeric probabilities replaced by qualitative constraints (Wellman 1990, Druzdzel & Henrion 1993)
- Potential benefits of using hybrid BN/QPN in GenIE:
  - Avoid requirement to acquire all numerical probabilities needed for BN
  - Supports use of qualitative values as found in corpus

## Previous Work on QPNs

- QPNs were developed for use in decision support systems (Wellman 1990, Druzdzal & Henrion 1993)
- qualitative belief propagation algorithm based on formal qualitative relations among nodes
    - influence (S), additive synergy (Y), product synergy (X)
    - provides formal account of ‘explaining away’
  - used to generate explanation of change in belief addressed to *expert* audience
    - but not like how genetic counselor talks to client!  
See comparison in our paper ..

## Current Work

- Analysis of canonical qualitative constraints for clinical genetics, e.g.
  - **negative product synergy** (like noisy-or):  
 $X^-$  ({OI-germline, OI-inherited}, OI-symptoms)
  - **positive additive synergy** (not like noisy-or):  
 $Y^+$  (CF/mother, CF/father, CF/child)  
where CF is Cystic Fibrosis (autosomal recessive)
  - **Enables:**  $X^+$  ({BRCA1, mutation}, BC)
  - **Inhibits:**  $Y^-$  ({masectomy, BRCA1}, BC)
- Defining discourse generation strategies using hybrid BN/QPN for KB

## **Conclusions and On-going Work**

- Effective patient communication in clinical genetics must address uncertainty
- Corpus study to understand how uncertainty used currently in patient communication
  - characterization in terms of form and function
  - to inform development of communicative strategies for use by GenIE
- Investigating hybrid BN/QPN to support GenIE's model of domain of discourse