From Diagnostics to Risk Assessment, Causality and Therapeutic intervention

Andrew Lynn
School of Computational and Integrative Sciences
Jawaharlal Nehru University

Current Diagnostics are threshold classifiers or images that require visual interpretation

Threshold based diagnostics

Biochemical assays

Microbiological assays

immunoblot analysis

Disease detection for Diabetes, Liver anomalies, Kidney issues etc.

Visual inspection based diagnostics

Histopathological

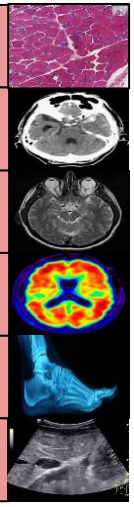
CT-scan

MRI

PET

Radiology

Ultrasound



Detection of tumors, artifacts and inflammations.

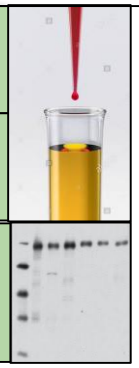
Machine Learning can be used to identify new multivariate diagnostics and automate visual interpretation

Threshold based diagnostics

Biochemical assays

Microbiological assays

immunoblot analysis



Data Mining and Machine learning

Discovery of new and improved multivariate diagnostics

Visual inspection based diagnostics

Histopathological

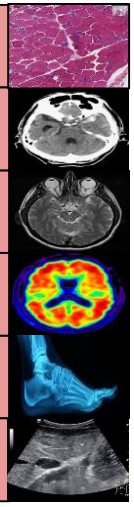
CT-scan

MRI

PET

Radiology

Ultrasound



Deep learning

Automate Visual Interpretation

Machine Learning can be used to identify new multivariate diagnostics and automate visual interpretation

Threshold based diagnostics

- Biochemical assays
- Microbiologic assays
- immunoblot analysis

Visual inspection based diagnostics

- Histopathologic
- CT-scar
- MRI
- PET
- Radiolog
- Ultrasou

REPORT

Detection and localization of surgically resectable cancers with a multi-analyte blood test

Joshua D. Cohen[1,2,3,4,5], Lu Li[6], Yuxuan Wang[1,2,3,4], Christopher Thoburn[3], Bahman Afsari[7], Ludmila Danilova[7], Christoph...
+ See all authors and affiliations

Science 23 Feb 2018:
Vol. 359, Issue 6378, pp. 926-930
DOI: 10.1126/science.aar3247

Article | Figures & Data | Info & Metrics | eLetters | PDF

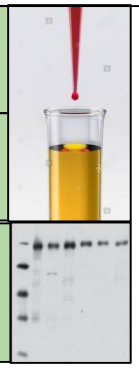
SEEK and you may find cancer earlier

Many cancers can be cured by surgery and/or systemic therapies when detected before they have metastasized. This clinical reality, coupled with the growing appreciation that cancer's rapid genetic evolution limits its response to drugs, have fueled interest in methodologies for earlier detection of the disease. Cohen *et al.* developed a noninvasive blood test, called CancerSEEK that can detect eight common human cancer types (see the Perspective by Kalinich and Haber). The test assesses eight circulating protein biomarkers and tumor-specific mutations in circulating DNA. In a study of 1000 patients previously diagnosed with cancer and 850 healthy control individuals, CancerSEEK detected cancer with a sensitivity of 69 to 98% (depending on cancer type) and 99% specificity.

Machine Learning can be used to identify new multivariate diagnostics and automate visual interpretation

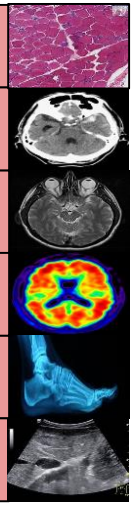
Threshold based diagnostics

- Biochemical assays
- Microbiological assays
- immunoblot analysis

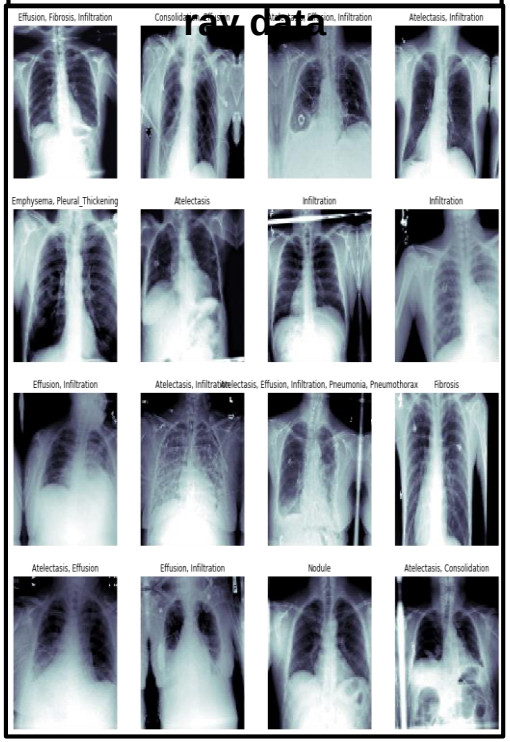


Visual inspection based diagnostics

- Histopathological
- CT-scan
- MRI
- PET
- Radiology
- Ultrasound

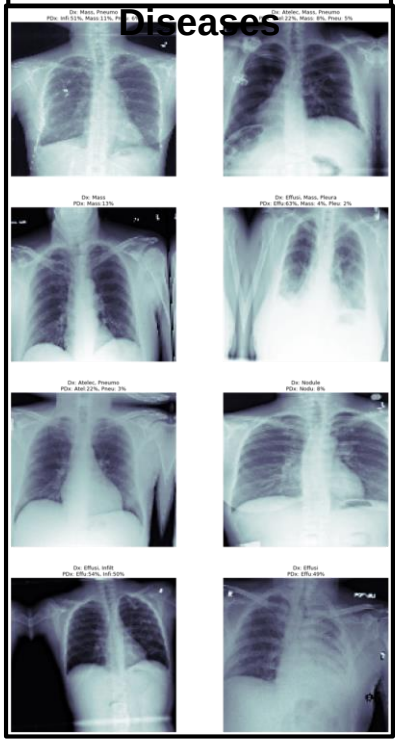


Input: Lung Diseases X-ray data



Deep learning

Output: Predicted Diseases



Automate Visual Interpretation

Genetic testing is used for risk prediction

Threshold based diagnostics

Biochemical assays

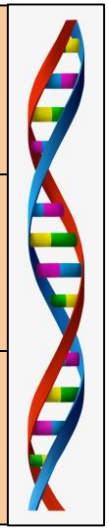
Microbiological assays

Genetics based diagnostics

Genome variation
SNP,Indel and Structural variation

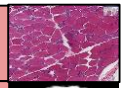
Epigenetics variations
Histone modifications and methylation

RNAseq
Transcript expression DEGs



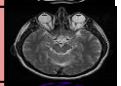
Visual inspection based diagnostics

Histopathological

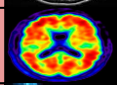


CT-scan

MRI



PET



Radiology

Ultrasound

Disease Detection, Risk Prediction, early stage disease detection
Eg; Breast cancer gene panel test, Thalassemia mutation test, Deletion studies in DMD

CURRENT DIAGNOSTICS

Data Mining and Machine learning

New Dimensions of genetic diagnostics tests and Better Risk prediction

FUTURE DIAGNOSTICS

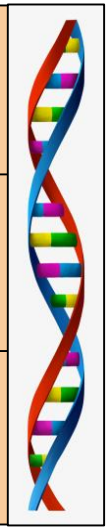
Genetic testing is used for risk prediction

Threshold based diagnostics

- **Biochemical assays**
- **Microbiological assays**

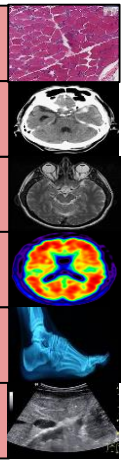
Genetics based diagnostics

- **Genome variation** *SNP,Indel and Structural variation*
- **Epigenetics variations** *Histone modifications and methylation*
- **RNAseq** *Transcript expression DEGs*



Visual inspection based diagnostics

- **Histopathological**
- **CT-scan**
- **MRI**
- **PET**
- **Radiology**
- **Ultrasound**

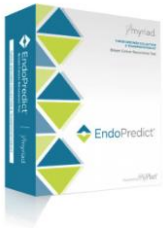


Disease Detection, Risk Prediction, early stage disease detection
Eg; Breast cancer gene panel test, Thalassemia mutation test, Deletion studies in DMD

BRACAnalysis

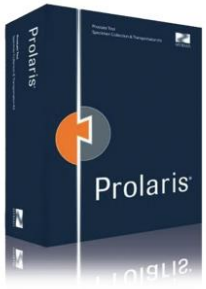
Test your risks for breast and ovarian cancers with BRACAnalysis. The innovator's BRCA test. For breast/ovarian cancer and family members.

EndoPredict



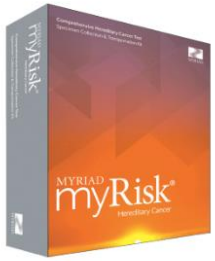
Decide if you can safely avoid chemotherapy and how beneficial chemotherapy would be. For breast cancer.

Prolaris



Guideline recommended test for treatment planning. For prostate cancer.

MyRisk



Test for your risk to cancer and get medical management recommendations. For individuals/family history of cancer.

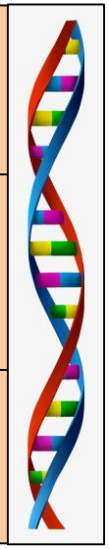
Novel diagnostics and risk predictors can be identified from modern DNA/RNA sequencing methods

Threshold based diagnostics

- **Biochemical assays**
- **Microbiological assays**

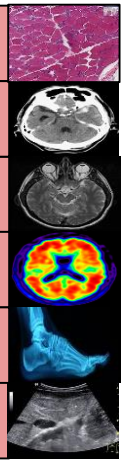
Genetics based diagnostics

- **Genome variation** *SNP,Indel and Structural variation*
- **Epigenetics variations** *Histone modifications and methylation*
- **RNAseq** *Transcript expression DEGs*



Visual inspection based diagnostics

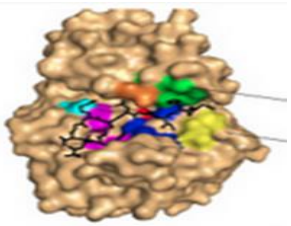
- **Histopathological**
- **CT-scan**
- **MRI**
- **PET**
- **Radiology**
- **Ultrasound**

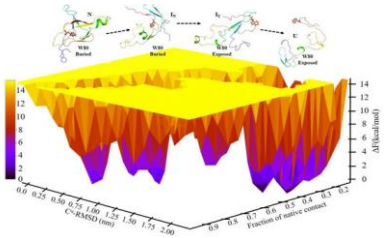


Epigenetic changes like CpG island hypomethylation or hypermethylation can be used as potential biomarkers in risk stratification and therapeutic response generation for various diseases like colorectal cancer, glioblastoma, breast cancer, asthma, neurodegenerative diseases etc.

Detection of relative abundance of transcript copies with tissue-specific effects is done commonly for cancer diagnosis and can also be used for detecting differentially expressed variants.

The impact of genomic variations can be understood from protein structure, dynamics and network analysis





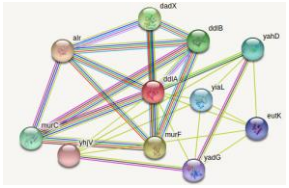
Protein 3D structure (PDB)

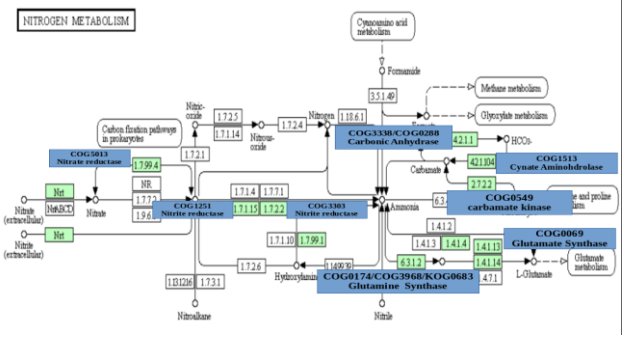
Understanding of protein's function and its role in disease

Genetic variations (SNPs, Indel and structural variations)

Network based modelling (STRING, Kegg Pathways)

Similarities and differences among distinct tumors





Threshold based diagnostics

Biochemical assays

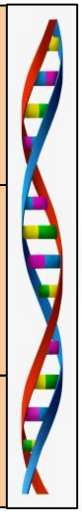
Microbiological assays

Genetics based diagnostics

Genome variation *SNP,Indel and Structural variation*

Epigenetics variations *Histone modifications and methylation*

RNAseq *Transcript expression DEGs*



Visual inspection based diagnostics

Histopathological

CT-scan

MRI

PET

Radiology

Ultrasound

Metabolic networks

Transcription networks

Protein Structure

Protein Protein Interaction

Signal transduction networks

Inter cellular signalling networks

Protein Dynamics

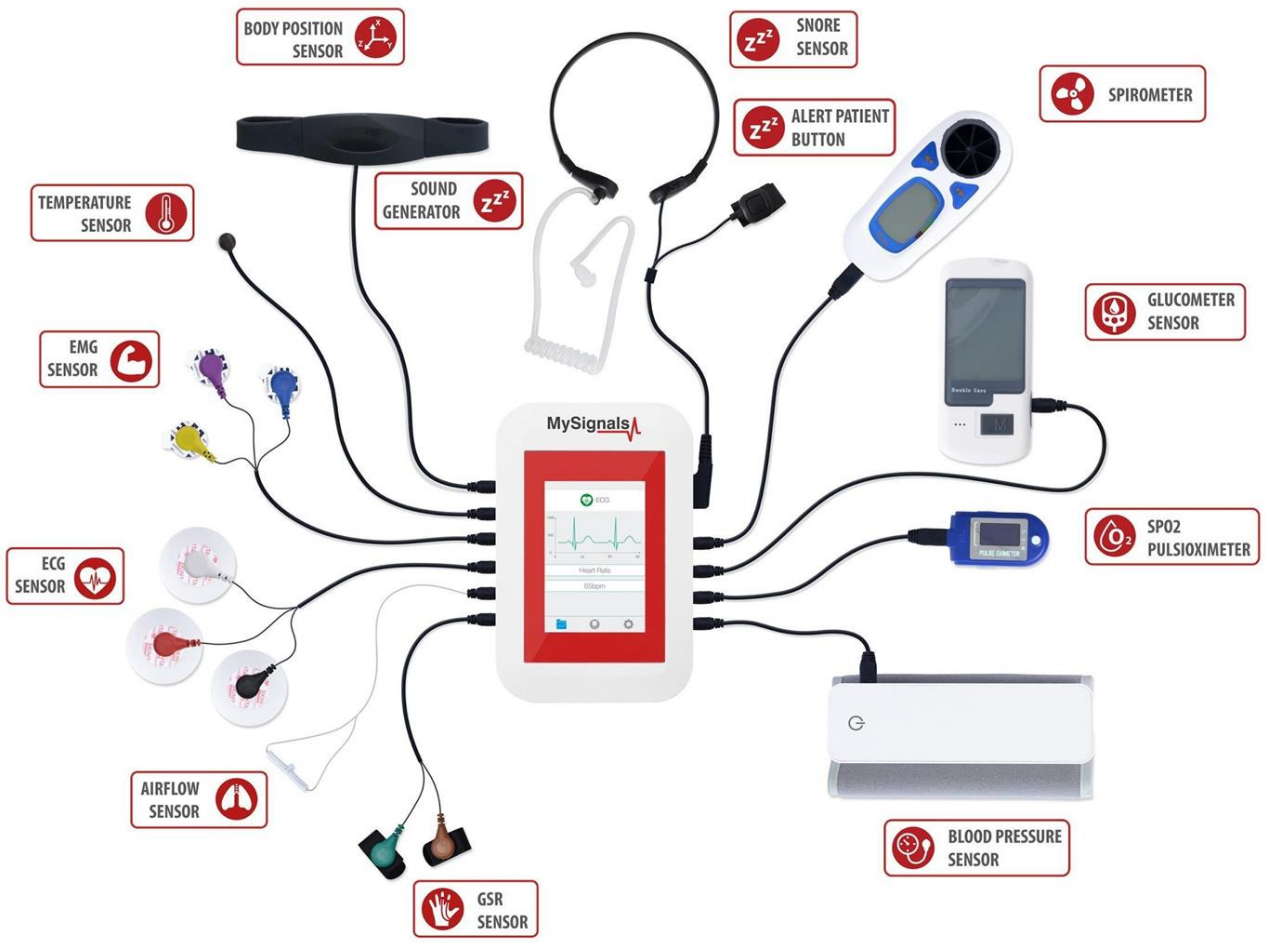
Biochemistry and Bioinformatics approaches

Causality

Therapeutic interventions

A similar impact can be studied with environmental factors using both data mining with feature extraction and network analysis

- Dietary habits
- Activity Measurement
- Fitness parameters
- Stress Levels
- Environmental conditions



A similar impact can be studied with environmental factors using both data mining with feature extraction and network analysis

Dietary habits

Activity Measurement

Fitness parameters

Stress Levels

Environmental conditions

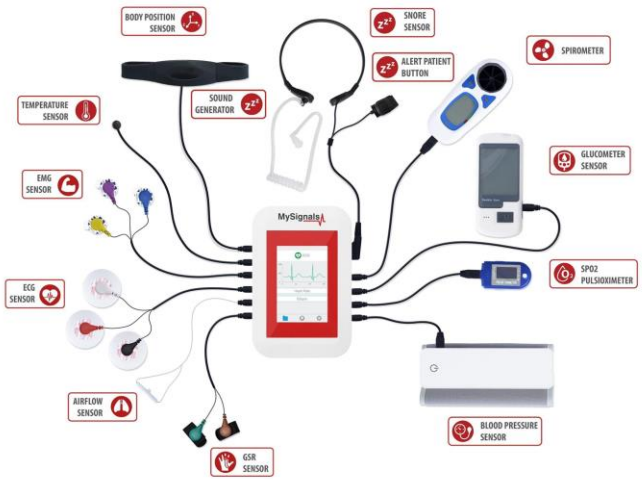
Data Mining and Machine learning

Network Analysis

Relation between Environmental factors, lifestyle and Risk assessment

Interconnection and Pathways informations

Systemic Cure



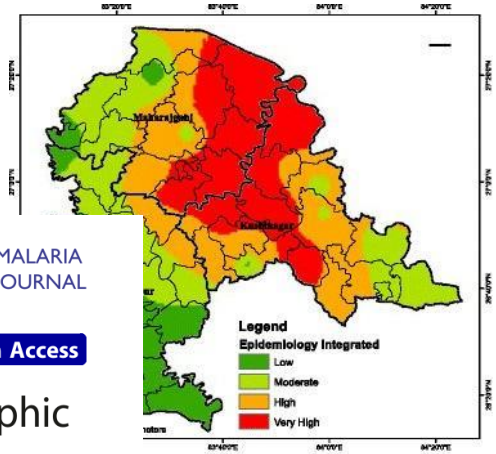
Qayum *et al. Malaria Journal* (2015) 14:192
DOI 10.1186/s12936-015-0685-4

MALARIA JOURNAL

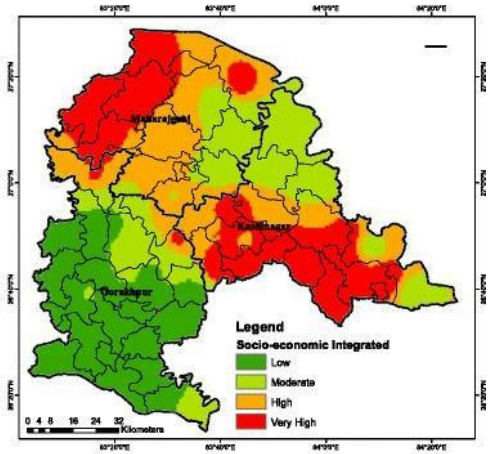
RESEARCH Open Access

Socio-economic, epidemiological and geographic features based on GIS-integrated mapping to identify malarial hotspots

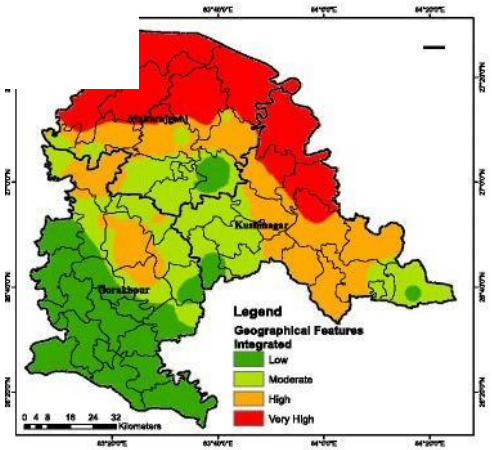
Abdul Qayum[1,2], Rakesh Arya[3], Pawan Kumar[4*] and Andrew M Lynn[1]



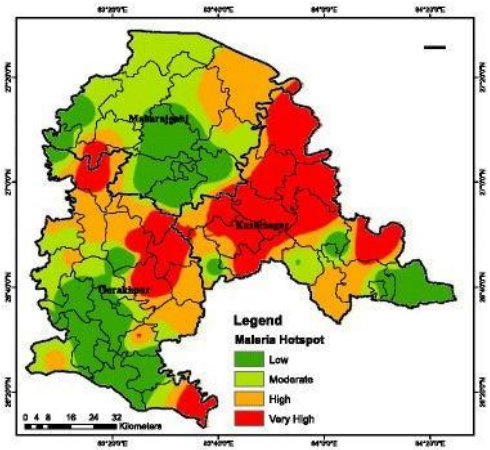
GIS integrated Epidemiology



9.2: GIS integrated Socio-economic



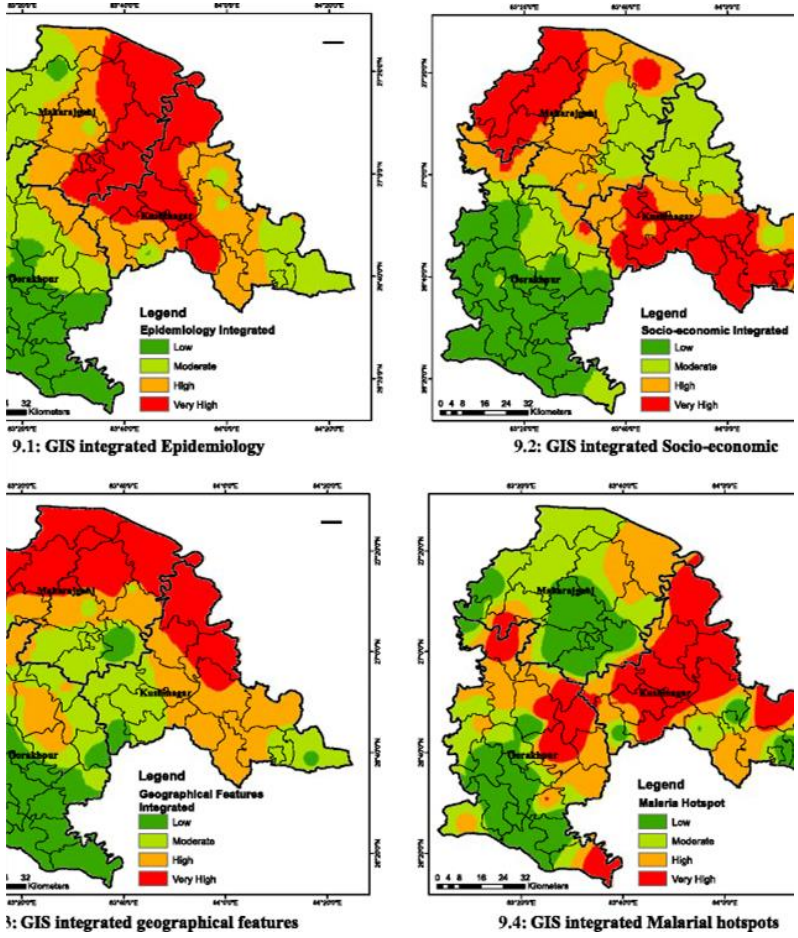
9.3: GIS integrated geographical features



9.4: GIS integrated Malarial hotspots

Integrated multiparameter data model for prediction of malaria hotspots

- 1. **Socio-economic aspects**
  - Population, child population, literacy, and workforce participation
- 2. **Epidemiology**
  - Annual parasitic index (API) and slides collected and examined)
- 3. **Geographical features**
  - Settlement, forest cover, water bodies, rainfall, relative humidity, and temperature



9.1: GIS integrated Epidemiology

9.2: GIS integrated Socio-economic

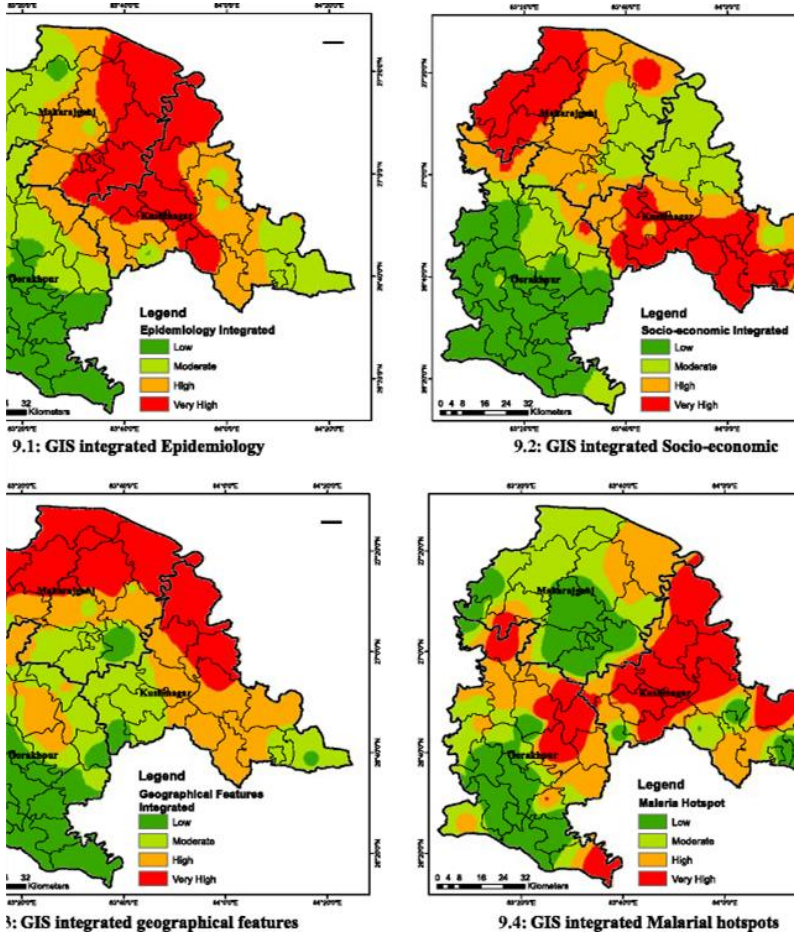
: GIS integrated geographical features

9.4: GIS integrated Malarial hotspots

- ***Malaria Journal, 2014***

Integrated multiparameter data model for prediction of malaria hotspots

- 1. **Socio-economic aspects**
  - Population, child population, literacy, and workforce participation
- 2. **Epidemiology**
  - Annual parasitic index (API) and slides collected and examined)
- 3. **Geographical features**
  - Settlement, forest cover, water bodies, rainfall, relative humidity, and temperature



9.1: GIS integrated Epidemiology

9.2: GIS integrated Socio-economic

: GIS integrated geographical features

9.4: GIS integrated Malarial hotspots

- ***Malaria Journal, 2014***

Threshold based diagnostics

Biochemical assays

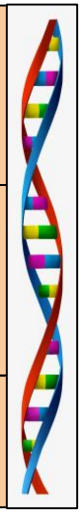
Microbiological assays

Genetics based diagnostics

Genome variation
SNP,Indel and Structural variation

Epigenetics variations
Histone modifications and methylation

RNAseq
Transcript expression DEGs



Visual inspection based diagnostics

Histopathological

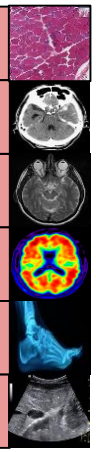
CT-scan

MRI

PET

Radiology

Ultrasound



Protein Structure

Metabolic networks

Transcription networks

Protein Protein Interaction

Signal transduction networks

Inter cellular signalling networks

Protein Dynamics

Biochemistry and Bioinformatics approaches

Causality

Therapeutic interventions

Dietary habits

Activity Measurement

Fitness parameters

Stress Levels

...

Environmental conditions