

Machine Learning in Toxicology: Towards Intelligent Access to the Content of Research

SACATM Meeting, September 2-3, 2020

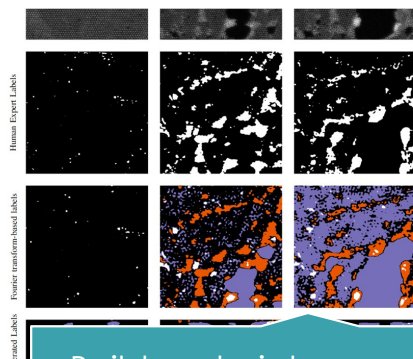
Robert M. Patton, Ph.D.

Computer Science and Mathematics Division

Oak Ridge National Laboratory

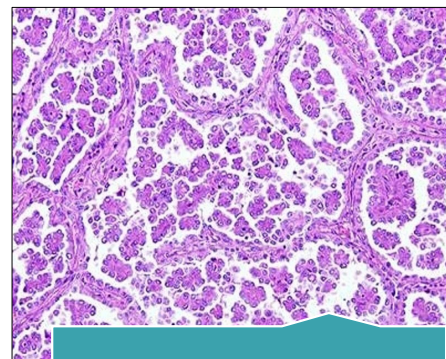
Computational Data Analytics

Core focus: developing innovative analytical methods for understanding large and complex datasets



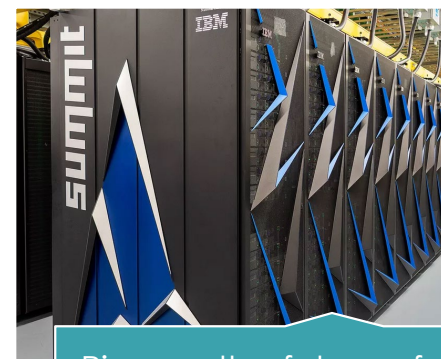
Build materials one atom as a time

Patton, R.M., Johnston, J.T., Young, S.R., Schuman, C.D., March, D.D., Potok, T.E., Rose, D.C., Lim, S.H., Karnowski, T.P., Ziatdinov, M.A. and Kalinin, S.V., **167-PFlops deep learning for electron microscopy: from learning physics to atomic manipulation.** *SC18* (p. 50). IEEE Press.



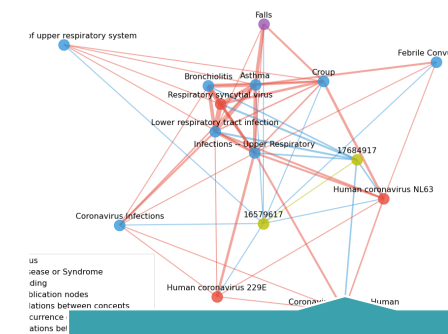
Fight cancer

Patton, R.M., Johnston, J.T., Young, S.R., Schuman, C.D., Potok, T.E., Rose, D.C., Lim, S.H., Chae, J., Hou, L., Abousamra, S. and Samaras, D., 2019. **Exascale Deep Learning to Accelerate Cancer Research.** *IEEE Big Data (2019)*



Pioneer the future of computing

T. Potok, C. Schuman, S. Young, R. Patton, F. Spedalieri, J. Liu, K.-T. Yao, G.S. Rose, and G. Chakma, **“A Study of Complex Deep Learning Networks on High Performance, Neuromorphic, and Quantum Computers,”** *ACM Journal of Emerging Technologies in Computing Systems*, vol. 14, no. 2, July 2018.



Accelerate biomedical discovery

R. Kannan, P. Sao, H. Lu, D. Herrmannova, R. Patton, T. Potok, V. Thakkar, R. Vuduc, **“Scalable Knowledge-Graph Analytics at 136 Petaflops/s,”** *Gordon Bell Submission, Under Review*, April 2020.

Supporting systematic review data extraction



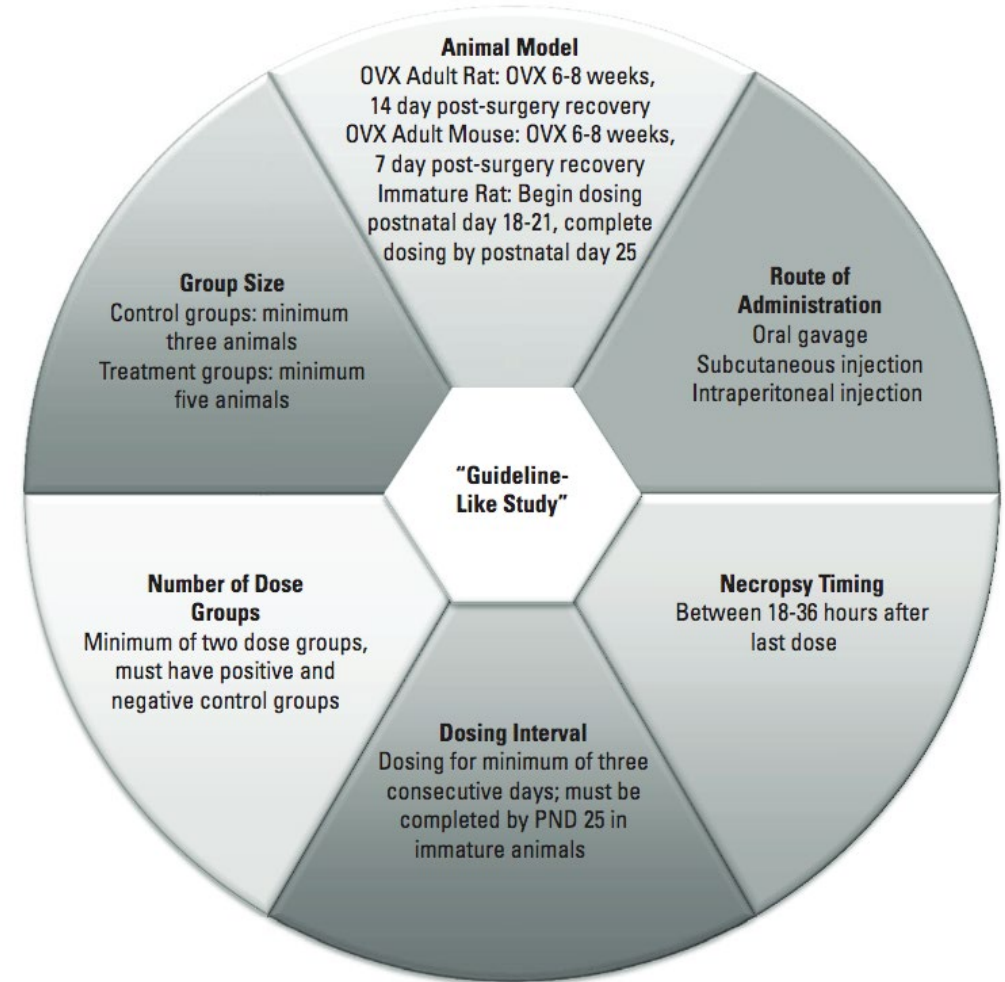
The Data Being Extracted

- OECD Test Guideline 440: Uterotrophic Bioassay in Rodents
 - The guideline consists of **6 minimal criteria (MC)**
 - **All six MC have to be met** for a study to be **guideline-like (GL)**
 - Below: manually annotated (GL) abstract showing the six MC

The intact female weanling version in the Organization for Economic Cooperation and Development (OECD) uterotrophic assay Test Guideline (TG) 440 is proposed as an alternative to the adult ovariectomized female version, because it does not involve surgical intervention (vs the ovariectomized version) and detects direct/indirect-acting estrogenic/anti-estrogenic substances (vs the ovariectomized version which detects only ^{GS} direct-acting estrogenic/antiestrogenic substances binding to the estrogen receptor). This validation study followed OECD TG 440, with ^{GS} six female weanling ^{GS} rats (postnatal day 21) ^{group size (GS)} per dose group and ^{no. of dose groups} six treatment groups. Females were weighed and dosed once daily by ^{oral gavage} admin. route for ^{dosing interval} three consecutive days, with one of six doses of 17 α -ethinyl estradiol in corn oil at 5 ml kg⁻¹ at ^{dose groups (neg. and pos. control)} 0 and 0.1–10 μ g kg⁻¹ per day. On ^{necropsy timing} postnatal day 24, the juvenile females were euthanized by CO₂ asphyxiation, weighed, livers weighed and uteri weighed wet and blotted. The presence or absence of vaginal patency was recorded. Absolute and relative (to terminal body weight) uterine wet and blotted weights and uterine luminal fluid weights were significantly increased at 3.0 and 10.0 (both P < 0.01) μ g kg⁻¹ per day, and increased to ~140% of control values at 1.0 μ g kg⁻¹ per day (not statistically significantly). In vivo body weights, weight changes, feed consumption, liver weights and terminal body weights were unaffected. Vaginal patency was not acquired in any female at any dose, although vaginal puckering was observed in one female at 10.0 μ g kg⁻¹ per day. Therefore, this intact weanling uterotrophic assay is validated in our laboratory for use under US and European endocrine toxicity testing programs/legislation.

OECD TG 440: Minimum criteria (MC)

- MC 1: Animal model
- MC 2: Group size
- MC 3: Route of administration
- MC 4: Number of dose groups
- MC 5: Dosing interval
- MC 6: Necropsy timing



Source: Kleinstreuer et al. (2016). *A Curated Database of Rodent Uterotrophic Bioactivity*.

Human Learning vs. Machine Learning

- A human non-expert can **correctly annotate criteria** given just a **description of the target information**

The intact female weanling version in the Organization for Economic Cooperation and Development (OECD) uterotrophic assay Test Guideline (TG) 440 is proposed as an alternative to the adult ovariectomized female version, because it does not involve surgical intervention (vs the ovariectomized version) and detects direct/indirect-acting estrogenic/anti-estrogenic substances (vs the ovariectomized version which detects only ^{GS} direct-acting estrogenic/antiestrogenic substances binding to the estrogen receptor). This validation study followed OECD TG 440, with ^{GS} six female weanling ^{GS} rats (postnatal day 21) ^{group size (GS)} per dose group and ^{no. of dose groups} six treatment groups. Females were weighed and dosed once daily by ^{admin. route} oral gavage for ^{dosing interval} three consecutive days, with one of six doses of 17 α -ethinyl estradiol in corn oil at 5 ml kg⁻¹ at ^{dose groups (neg. and pos. control)} 0 and 0.1–10 μ g kg⁻¹ per day. On ^{necropsy timing} postnatal day 24, the juvenile females were euthanized by CO₂ asphyxiation, weighed, livers weighed and uteri weighed wet and blotted. The presence or absence of vaginal patency was recorded. Absolute and relative (to terminal body weight) uterine wet and blotted weights and uterine luminal fluid weights were significantly increased at 3.0 and 10.0 (both P < 0.01) μ g kg⁻¹ per day, and increased to ~140% of control values at 1.0 μ g kg⁻¹ per day (not statistically significantly). In vivo body weights, weight changes, feed consumption, liver weights and terminal body weights were unaffected. Vaginal patency was not acquired in any female at any dose, although vaginal puckering was observed in one female at 10.0 μ g kg⁻¹ per day. Therefore, the intact female weanling uterotrophic assay is validated in our laboratory for use under US and European endocrine toxicity testing programs/legislation.

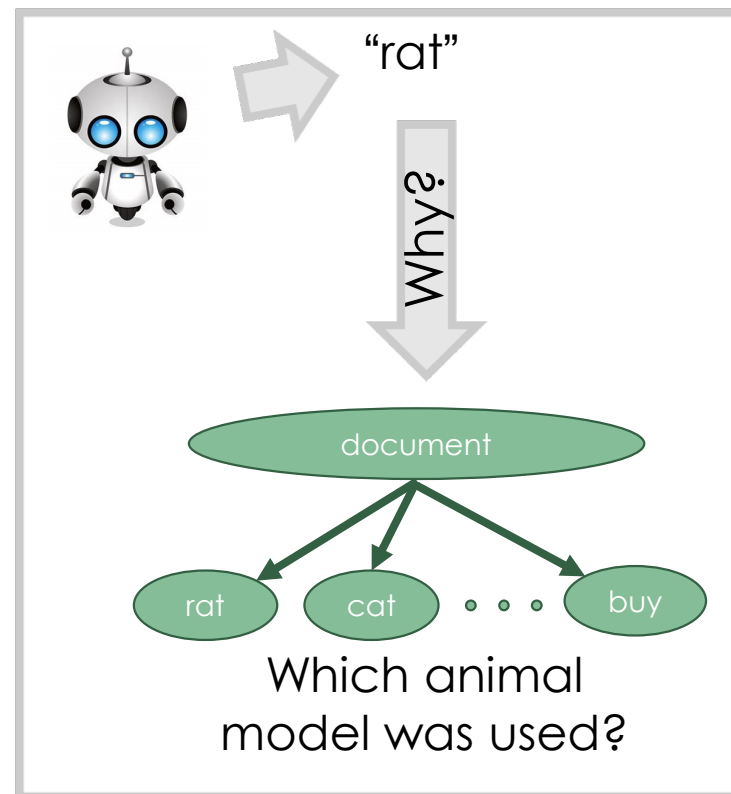
Our approach

- Goal: **extract information for the purpose of identifying GL/non-GL documents**

- Two approaches:

1. **Classify, then extract:** Can we train a good classifier? If so, extract learned patterns (fully supervised)

2. **Extract, then classify:** Extract text segments relevant to criteria descriptions (unsupervised), then classify the extracted segments



The circular diagram is divided into six segments, each representing a criterion for a 'Guideline-Like Study':

- Animal Model:** OVX Adult Rat: OVX 6-8 weeks, 14 day post-surgery recovery; OVX Adult Mouse: OVX 6-8 weeks, 7 day post-surgery recovery; Immature Rat: Begin dosing postnatal day 18-21, complete dosing by postnatal day 25.
- Route of Administration:** Oral gavage; Subcutaneous injection; Intraperitoneal injection.
- Necropsy Timing:** Between 18-36 hours after last dose.
- Dosing Interval:** Dosing for minimum of three consecutive days; must be completed by PND 25 in immature animals.
- Number of Dose Groups:** Minimum of two dose groups; must have positive and negative control groups.
- Group Size:** Control groups: minimum three animals; Treatment groups: minimum five animals.

Sentence 1: Pre-pubertal (day 18 of life) female mice were randomized to receive placebo, 200 mg/kg CTX, or 120 mg/kg CTX. **Similarity to MC1: 0.5478** (Green arrow and checkmark)

Sentence 2: The dosages of CTX used were based on previous studies, which demonstrated a significant dose-dependent ovarian toxicity. **Similarity to MC1: 0.3887** (Red arrow)

Dataset

1. Unstructured text of criteria descriptions (MC 1 – 6)
2. **670 research publications** with results for 2,615 uterotrophic bioassays
 - *A curated database of rodent uterotrophic bioactivity. (2015). Kleinstreuer et. al. Environmental health perspectives.*
 - **~120 publications** (~18%) contain GL studies (~540 out of 2,615)

Example results

Example 1: Necropsy timing (correct answer underlined)

The intact female weanling version in the Organization for Economic Cooperation and Development (OECD) uterotrophic assay Test Guideline (TG) 440 is proposed as an alternative to the adult ovariectomized female version, because it does not involve surgical intervention (vs the ovariectomized version) and detects direct/indirect-acting estrogenic/anti-estrogenic substances (vs the ovariectomized version which detects only direct-acting estrogenic/anti-estrogenic substances binding to the estrogen receptor). This validation study followed OECD TG 440, with six female weanling rats (postnatal day 21) per dose group and six treatment groups. Females were weighed and dosed once daily by oral gavage for three consecutive days, with one of six doses of 17alpha-ethinyl estradiol in corn oil at 5 ml kg⁻¹ at 0 and 0.1-10 microg kg⁻¹ per day. On postnatal day 24, the juvenile females were euthanized by CO(2) asphyxiation, weighed, livers weighed and uteri weighed wet and blotted. The presence or absence of vaginal patency was recorded. Absolute and relative (to terminal body weight) uterine wet and blotted weights and uterine luminal fluid weights were significantly increased at 3.0 and 10.0 (both P < 0.01) microg kg⁻¹ per day, and increased to ~140% of control values at 1.0 microg kg⁻¹ per day (not statistically significantly). In vivo body weights, weight changes, feed consumption, liver weights and terminal body weights were unaffected. Vaginal patency was not acquired in any female at any dose, although vaginal puckering was observed in one female at 10.0 microg kg⁻¹ per day. Therefore, this intact weanling uterotrophic assay is validated in our laboratory for use under US and European endocrine toxicity testing programs/legislation.

Yellow: sentence most similar to description
Red: segment most similar to description

Example 2: Number of dose groups (correct answer underlined)

The intact female weanling version in the Organization for Economic Cooperation and Development (OECD) uterotrophic assay Test Guideline (TG) 440 is proposed as an alternative to the adult ovariectomized female version, because it does not involve surgical intervention (vs the ovariectomized version) and detects direct/indirect-acting estrogenic/anti-estrogenic substances (vs the ovariectomized version which detects only direct-acting estrogenic/anti-estrogenic substances binding to the estrogen receptor). This validation study followed OECD TG 440, with six female weanling rats (postnatal day 21) per dose group and six treatment groups. Females were weighed and dosed once daily by oral gavage for three consecutive days, with one of six doses of 17alpha-ethinyl estradiol in corn oil at 5 ml kg⁻¹ at 0 and 0.1-10 microg kg⁻¹ per day. On postnatal day 24, the juvenile females were euthanized by CO(2) asphyxiation, weighed, livers weighed and uteri weighed wet and blotted. The presence or absence of vaginal patency was recorded. Absolute and relative (to terminal body weight) uterine wet and blotted weights and uterine luminal fluid weights were significantly increased at 3.0 and 10.0 (both P < 0.01) microg kg⁻¹ per day, and increased to ~140% of control values at 1.0 microg kg⁻¹ per day (not statistically significantly). In vivo body weights, weight changes, feed consumption, liver weights and terminal body weights were unaffected. Vaginal patency was not acquired in any female at any dose, although vaginal puckering was observed in one female at 10.0 microg kg⁻¹ per day. Therefore, this intact weanling uterotrophic assay is validated in our laboratory for use under US and European endocrine toxicity testing programs/legislation.

Next steps

- Move beyond solely identifying occurrence of target terms in text by **incorporating context**
 - For example, distinguish cases where a chemical used as experimental treatment or as part of a laboratory procedure
- Jin et al. demonstrated biomedical contextual word vectors capable of **distinguishing different meanings of the same word**

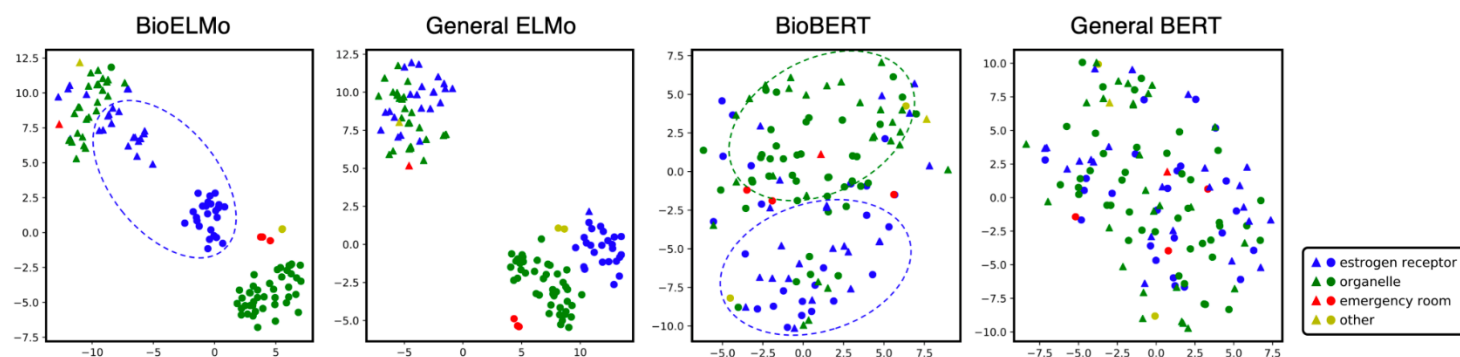
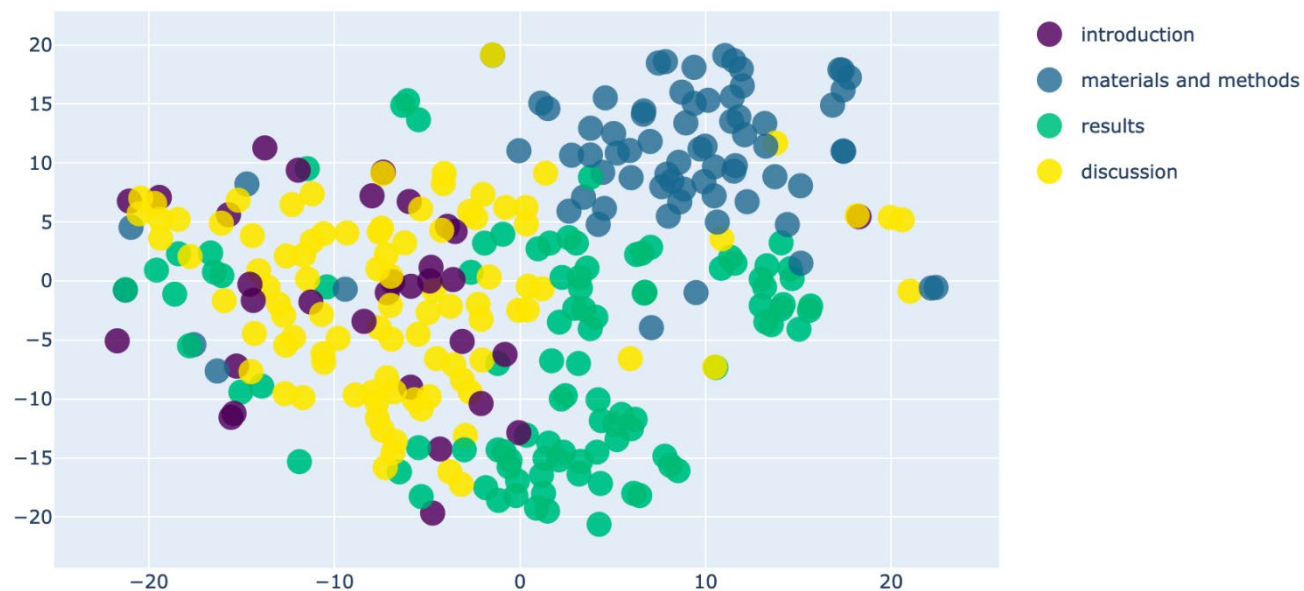


Figure 3: t-SNE visualizations of the token **ER** embeddings in different contexts by BioELMo, general ELMo, BioBERT and general BERT. ● and ▲ represent **ER** mentions within and outside of parentheses, respectively. Colors refer to different actual meanings of the **ER** mention.

Source: Jin et al. (2019). Probing Biomedical Embeddings from Language Models. 3rd Workshop on Evaluating Vector Space Representations for NLP, NAACL 2019.

Next steps

- BioELMo word vectors for all occurrences of the word "rat" in 50 uterotrophic bioassay publications
 - Color represents paper section
- Next steps: leverage contextual vectors to **enable fast identification of relevant mentions**



Current limitations and future work

- Problems:

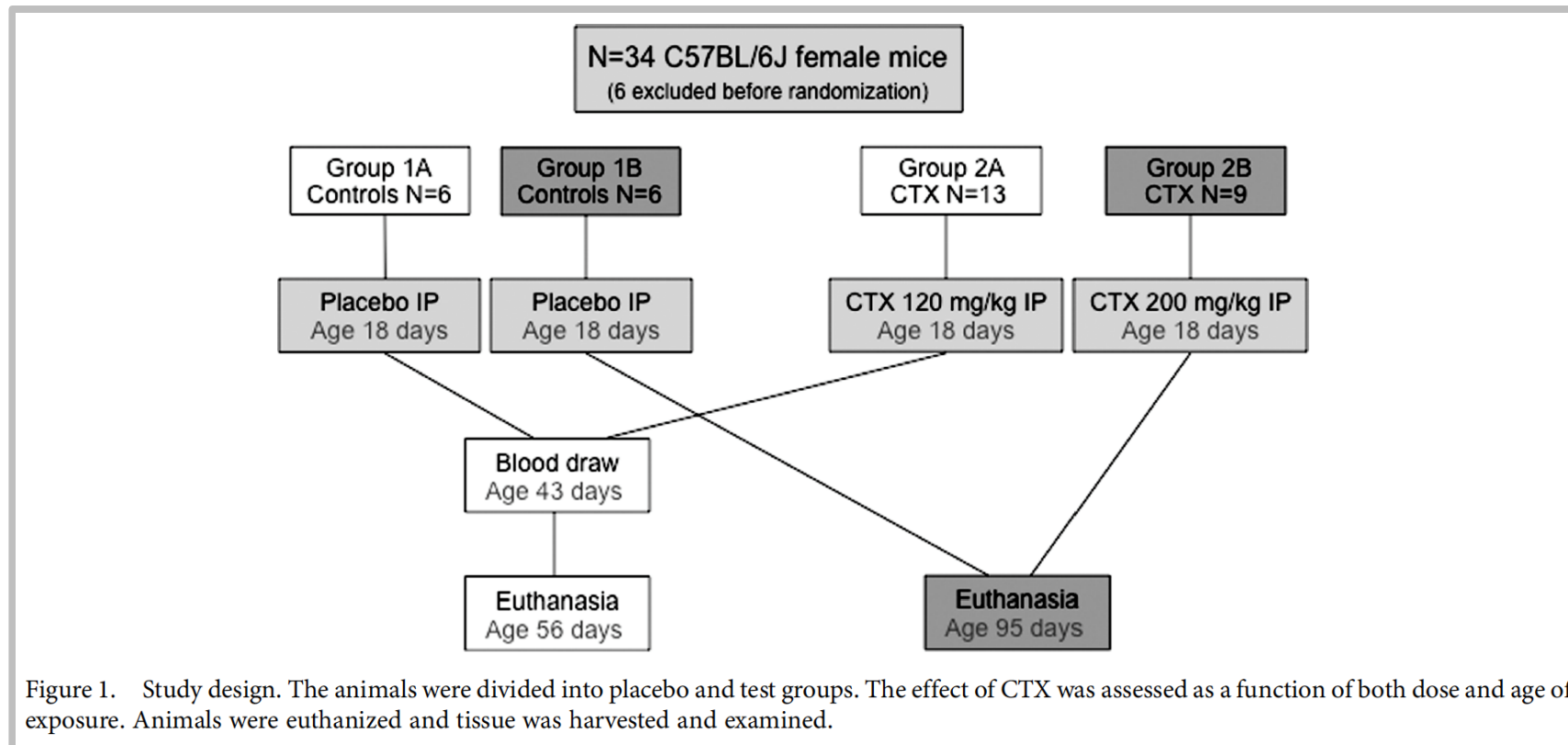
- “Semantic implication”
- Tables/figures
- Multiple studies
- Answer not in text

- Improvements

- Making extraction more accurate
 - Applying robust evaluation metrics toward the extracted text
 - Combining both (supervised and unsupervised) approaches
- Labeling extracted text
- Using ensembles to improve classification / joint modelling of MC labels (e.g. using Deep Learning)

Current limitations

- Tables/figures



Extracting structured data from scientific documents



Long term goal

Nurkiewicz et al.

intubated to ensure a patent airway, and the right carotid artery was cannulated to measure arterial pressure. The right spinotrapezius muscle was then extirpated, superfused with an electrolyte solution, and prepared for microscopic observation as previously described (Nurkiewicz et al. 2004).

The animal preparation was then transferred to the stage of an intravital microscope. Video images were displayed and videotaped for off-line analysis. During videotape replay, arteriolar inner diameters were measured and venular leukocyte adhesion was quantified.

Experimental protocols. Protocol 1. Arteriolar endothelium-dependent dilation was evaluated by assessing the capacity for Ca²⁺-dependent endothelial nitric oxide formation in response to intraluminal infusion of the calcium ionophore A23187 (Sigma Chemical Co., St. Louis, MO). Glass micropipettes were filled with a 10⁻⁷ M solution of A23187 and inserted into the arteriolar lumen, and A23187

was then infused directly into the flow stream for 2-min periods at ejection pressures of 5, 10, 20, and 40 psi (Nurkiewicz et al. 2004). A 2-min recovery period followed each ejection. At the end of all intravital experiments, adenosine (ADO) was added to the superfusate (10⁻⁴ M final concentration) to fully dilate the microvascular network and determine the passive diameter of each arteriole studied.

Protocol 2. To evaluate arteriolar responsiveness to adrenergic stimulation, phenylephrine (PHE) was iontophoretically applied to individual arterioles in rats after exposure to either saline or 0.25 mg ROFA. Micropipettes were filled with a 50-mM solution of PHE in distilled water. The pipette tip was placed in light contact with the arteriolar wall, and a current programmer delivered continuous 2-min ejection currents of 50, 100, and 200 nA (randomly). A 2-min recovery period followed each application. To exclude the possibility that adrenergic stimulation could increase NO

production and therefore attenuate the observed constrictions, these experiments were performed during NO synthase (NOS) inhibition with N^G-monomethyl-L-arginine (10⁻⁴ M final superfusate concentration).

Protocol 3. Adhering or rolling leukocytes in first-order venules of rats after exposure to either saline, 0.25 mg ROFA, or 0.1 mg TiO₂ were quantified to characterize microvascular inflammation. Leukocytes that were either stationary or moving but in constant contact with the venular wall for at least 200 μm were counted for 1 min in each venule studied.

Protocol 4. Oxidant activity in the arteriolar wall was measured with the tetranitroblue tetrazolium (TNBT) reduction method, which provides a general index of microvascular oxidant stress (Lenda and Boegehold 2002). After 1 hr of continuous exposure to 2% TNBT superfusion, the spinotrapezius muscle was fixed with a 10% formalin solution and excised. The tissue was then viewed with bright-field microscopy, and images of microvessels were digitized and analyzed. Using a 1 × 5 μm photometric window, a series of average pixel intensity measurements were made along the vessel wall and in extravascular regions immediately adjacent to the wall. To assess microvascular wall levels of formazan (the reduction product of TNBT and therefore an index of oxidant activity), the measured pixel intensities were used to calculate microvascular wall light absorption (A): $A = \ln(I_0/I)$, where I_0 is the vessel intensity and I is the intensity for the adjacent extravascular region. The amount of formazan formed is proportional to the level of oxidant activity, and calculated light absorption is linearly related to the amount of formazan present (Lenda and Boegehold 2002).

Data and statistical analyses. Arteriolar diameter (D , in micrometers) was sampled at 10-sec intervals. Resting vascular tone was calculated for each vessel as follows: $\text{tone} = [(D_{\text{pass}} - D)/D_{\text{pass}}] \times 100$, where D_{pass} is

Table 1. Profiles of experimental animals used for intravital studies.

Experimental group	N	Age (days)	Weight (g)	Mean arterial pressure (mm Hg)
Saline control	15	55 ± 4	236 ± 4	98 ± 5
0.1 mg TiO ₂	5	50 ± 1	275 ± 9*	108 ± 6
0.1 mg ROFA	4	52 ± 1	225 ± 3	102 ± 9
0.25 mg TiO ₂	3	56 ± 3	219 ± 7	91 ± 15
0.25 mg ROFA	11	50 ± 2	221 ± 8	94 ± 7

N, number of rats. Values are mean ± SE.
*p < 0.05 compared with all other groups.

Table 2. BAL data from saline-treated and TiO₂- and ROFA-exposed rats.

Experimental group	Cellular content of PMNL (10 ⁶ cells/rat)	Albumin (mg/mL)	LDH (U/L)	Total AM CL
Saline control	0.93 ± 0.11	0.13 ± 0.02	59 ± 10	7.50 ± 1.63
0.1 mg TiO ₂	1.73 ± 0.36*	0.14 ± 0.04	67 ± 13	5.21 ± 1.60
0.1 mg ROFA	1.24 ± 0.28	0.23 ± 0.05	75 ± 2	14.47 ± 2.60***
0.25 mg TiO ₂	1.11 ± 0.15	0.19 ± 0.02	65 ± 8	3.64 ± 0.73
0.25 mg ROFA	1.91 ± 0.20*	0.17 ± 0.01	46 ± 4	17.42 ± 1.11***

N = 22 rats for saline control; N = 5–7 rats for all other doses. Values are mean ± SE. CL = counts per minute × 10³/0.25 × 10⁶ AM/15 min.
*p < 0.05 compared with saline. **p < 0.05 compared with 0.1 mg TiO₂ and p < 0.05 compared with 0.25 mg TiO₂.
***p < 0.05 compared with saline; similar results were obtained with TiO₂.

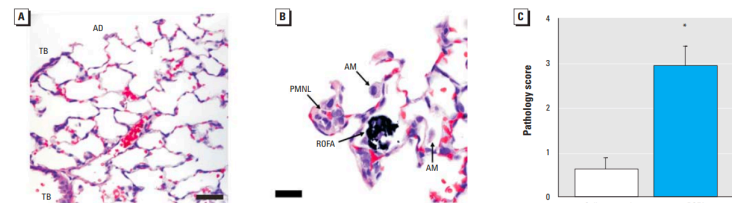


Figure 1. Histologic evidence of focal pulmonary alveolitis 24 hr after PM exposure. (A) and (B) are representative findings from five saline-treated rats and five rats exposed to 0.1 mg ROFA. (A) Saline control showing no morphologic alterations. Abbreviations: AD, alveolar duct; TB, terminal bronchioles. Bar = 50 μm. (B) Histopathologic alterations in a ROFA-exposed rat. Agglomerated ROFA particles in an alveolar space can be seen near an alveolar duct. The ROFA particles do not transmit light and therefore appear black when viewed in the light microscope. AMs are frequently observed in alveoli near ROFA particles and are intimately associated with the agglomerated ROFA. PMNLs are present in lesser numbers near ROFA particles, most frequently in the interstitium. Bar = 20 μm. (C) Mean alveolitis pathology scores.
*p < 0.05 compared with saline; similar results were obtained with TiO₂.

Table 1. Profiles of experimental animals used for intravital studies.

Experimental group	N	Age (days)	Weight (g)	Mean arterial pressure (mm Hg)
Saline control	15	55 ± 4	236 ± 4	98 ± 5
0.1 mg TiO ₂	5	50 ± 1	275 ± 9*	108 ± 6
0.1 mg ROFA	4	52 ± 1	225 ± 3	102 ± 9
0.25 mg TiO ₂	3	56 ± 3	219 ± 7	91 ± 15
0.25 mg ROFA	11	50 ± 2	221 ± 8	94 ± 7

N, number of rats. Values are mean ± SE.
*p < 0.05 compared with all other groups.

TiO₂ groups: Group size

Dose group	Group size
Control	15
0.1 mg	5
0.25 mg	3

TiO₂ groups: Age

Dose group	Age (days)
Control	Mean = 55, SE = 4
0.1 mg	Mean = 50, SE = 1
0.25 mg	Mean = 56, SE = 3

Short term goal

Nurkiewicz et al.

intubated to ensure a patent airway, and the right carotid artery was cannulated to measure arterial pressure. The right spinotrapezius muscle was then extirpated, superfused with an electrolyte solution, and prepared for microscopic observation as previously described (Nurkiewicz et al. 2004).

The animal preparation was then transferred to the stage of an intravital microscope. Video images were displayed and videotaped for off-line analysis. During videotape replay, arteriolar inner diameters were measured and venular leukocyte adhesion was quantified.

Experimental protocols. Protocol 1. Arteriolar endothelium-dependent dilation was evaluated by assessing the capacity for Ca²⁺-dependent endothelial nitric oxide formation in response to intraluminal infusion of the calcium ionophore A23187 (Sigma Chemical Co., St. Louis, MO). Glass micropipettes were filled with a 10⁻⁷ M solution of A23187 and inserted into the arteriolar lumen, and A23187

was then infused directly into the flow stream for 2-min periods at ejection pressures of 5, 10, 20, and 40 psi (Nurkiewicz et al. 2004). A 2-min recovery period followed each ejection. At the end of all intravital experiments, adenosine (ADO) was added to the superfusate (10⁻⁴ M final concentration) to fully dilate the microvascular network and determine the passive diameter of each arteriole studied.

Protocol 2. To evaluate arteriolar responsiveness to adrenergic stimulation, phenylephrine (PHE) was iontophoretically applied to individual arterioles in rats after exposure to either saline or 0.25 mg ROFA. Micropipettes were filled with a 50-mM solution of PHE in distilled water. The pipette tip was placed in light contact with the arteriolar wall, and a current programmer delivered continuous 2-min ejection currents of 50, 100, and 200 nA (randomly). A 2-min recovery period followed each application. To exclude the possibility that adrenergic stimulation could increase NO

production and therefore attenuate the observed constrictions, these experiments were performed during NO synthase (NOS) inhibition with N^G-monomethyl-L-arginine (10⁻⁴ M final superfusate concentration).

Protocol 3. Adhering or rolling leukocytes in first-order venules of rats after exposure to either saline, 0.25 mg ROFA, or 0.1 mg TiO₂ were quantified to characterize microvascular inflammation. Leukocytes that were either stationary or moving but in constant contact with the venular wall for at least 200 μm were counted for 1 min in each venule studied.

Protocol 4. Oxidant activity in the arteriolar wall was measured with the tetranitroblue tetrazolium (TNBT) reduction method, which provides a general index of microvascular oxidant stress (Lenda and Boegehold 2002). After 1 hr of continuous exposure to 2% TNBT superfusion, the spinotrapezius muscle was fixed with a 10% formalin solution and excised. The tissue was then viewed with bright-field microscopy, and images of microvessels were digitized and analyzed. Using a 1 × 5 μm photometric window, a series of average pixel intensity measurements were made along the vessel wall and in extravascular regions immediately adjacent to the wall. To assess microvascular wall levels of formazan (the reduction product of TNBT and therefore an index of oxidant activity), the measured pixel intensities were used to calculate microvascular wall light absorption (A): $A = \ln(I_0/I)$, where I_0 is the vessel intensity and I is the intensity of the adjacent extravascular region. The amount of formazan formed is proportional to the level of oxidant activity, and calculated light absorption is linearly related to the amount of formazan present (Lenda and Boegehold 2002).

Data and statistical analyses. Arteriolar diameter (D , in micrometers) was sampled at 10-sec intervals. Resting vascular tone was calculated for each vessel as follows: $\text{tone} = [(D_{\text{pass}} - D)/D_{\text{pass}}] \times 100$, where D_{pass} is

Table 1. Profiles of experimental animals used for intravital studies.

Experimental group	N	Age (days)	Weight (g)	Mean arterial pressure (mm Hg)
Saline control	15	55 ± 4	236 ± 4	98 ± 5
0.1 mg TiO ₂	5	50 ± 1	275 ± 9*	108 ± 6
0.1 mg ROFA	4	52 ± 1	225 ± 3	102 ± 9
0.25 mg TiO ₂	3	56 ± 3	219 ± 7	91 ± 15
0.25 mg ROFA	11	50 ± 2	221 ± 8	94 ± 7

N, number of rats. Values are mean ± SE.
*p < 0.05 compared with all other groups.

Table 2. BAL data from saline-treated and TiO₂- and ROFA-exposed rats.

Experimental group	Cellular content of PMNLs (10 ⁶ cells/rat)	Albumin (mg/mL)	LDH (U/L)	Total AM CL
Saline control	0.93 ± 0.11	0.13 ± 0.02	59 ± 10	7.50 ± 1.63
0.1 mg TiO ₂	1.73 ± 0.36*	0.14 ± 0.04	67 ± 13	5.21 ± 1.60
0.1 mg ROFA	1.24 ± 0.28	0.23 ± 0.05	75 ± 2	14.47 ± 2.60**
0.25 mg TiO ₂	1.11 ± 0.15	0.19 ± 0.02	65 ± 8	3.64 ± 0.73
0.25 mg ROFA	1.91 ± 0.20*	0.17 ± 0.01	46 ± 4	17.42 ± 1.11**

N = 22 rats for saline control; N = 5–7 rats for all other doses. Values are mean ± SE. CL = counts per minute × 10³/0.25 × 10⁶ AM/15 min.
*p < 0.05 compared with saline. **p < 0.05 compared with 0.1 mg TiO₂ and p < 0.05 compared with 0.25 mg TiO₂.

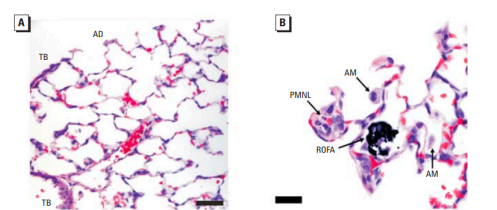


Figure 1. Histologic evidence of focal pulmonary alveolitis 24 hr after PM exposure. (A) and (B) are representative findings from five saline-treated rats and five rats exposed to 0.1 mg ROFA. (A) Saline control showing no morphologic alterations. Abbreviations: AD, alveolar duct; TB, terminal bronchioles. Bar = 50 μm. (B) Histopathologic alterations in a ROFA-exposed rat. Agglomerated ROFA particles in an alveolar space can be seen near an alveolar duct. The ROFA particles do not transmit light and therefore appear black when viewed in the light microscope. AMs are frequently observed in alveoli near ROFA particles and are intimately associated with the agglomerated ROFA. PMNLs are present in lesser numbers near ROFA particles, most frequently in the interstitium. Bar = 20 μm. (C) Mean alveolitis pathology scores.
*p < 0.05 compared with saline; similar results were obtained with TiO₂.

Table 1. Profiles of experimental animals used for intravital studies.

Experimental group	N	Age (days)	Weight (g)	Mean arterial pressure (mm Hg)
Saline control	15	55 ± 4	236 ± 4	98 ± 5
0.1 mg TiO ₂	5	50 ± 1	275 ± 9*	108 ± 6
0.1 mg ROFA	4	52 ± 1	225 ± 3	102 ± 9
0.25 mg TiO ₂	3	56 ± 3	219 ± 7	91 ± 15
0.25 mg ROFA	11	50 ± 2	221 ± 8	94 ± 7

N, number of rats. Values are mean ± SE.
*p < 0.05 compared with all other groups.

Caption: Table 1. Profiles of experimental animals ...

Experimental group	N	Age (days)	Weight (g)	...
Saline control	15	55 ± 4	236 ± 4	...
0.1 mg TiO ₂	5	50 ± 1	275 ± 9	...
0.1 mg ROFA	4	52 ± 1
0.25 mg TiO ₂	3
0.25 mg ROFA	11

Notes: N, number of rats. Values are mean ± SE. ...

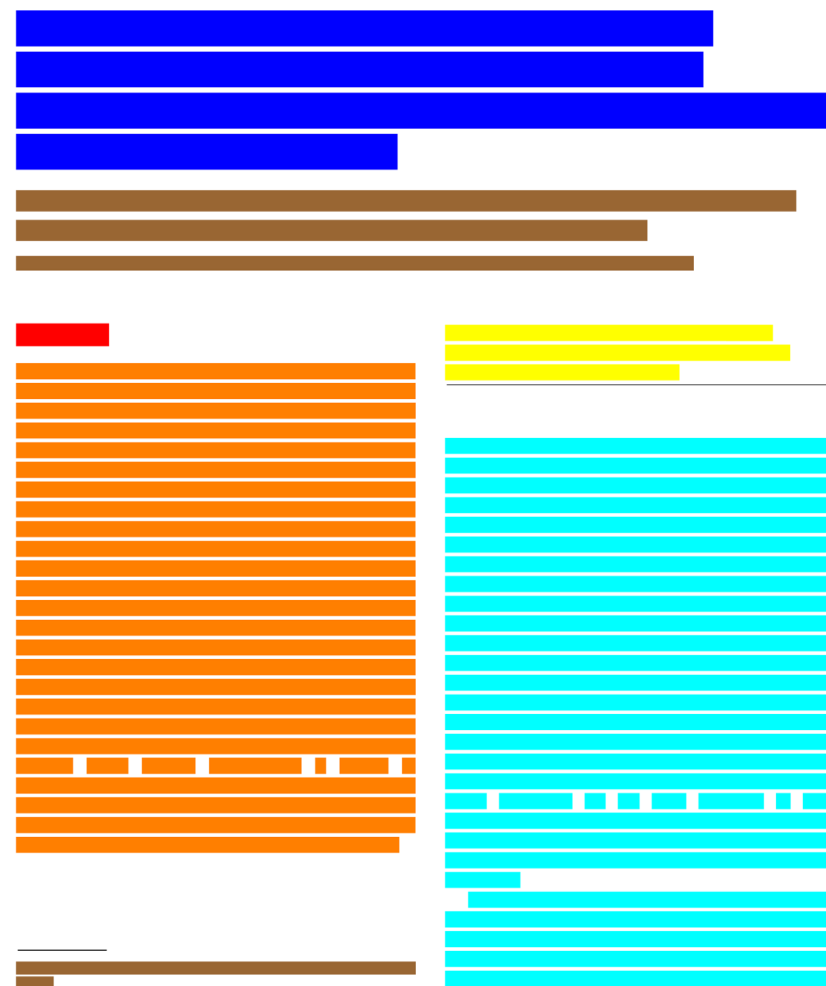
DeepPDF object detection

- Deep Neural Networks highly successful in image analysis
- Apply same principle to scientific documents
- Used network which was developed for medical image segmentation
 - Performs well on small datasets

Stahl et al. (2018) DeepPDF: A Deep Learning Approach to Analyzing PDFs. 2018 Workshop on Mining Scientific Publications at the Eleventh International Conference on Language Resources and Evaluation (LREC 2018).

Abdominal
Imaging

INVITED FEATURE SECTION



DeepPDF object detection

- Deep Neural Networks highly successful in image analysis
- Apply same principle to scientific documents
- Used network which was developed for medical image segmentation
 - Performs well on small datasets

Stahl et al. (2018) DeepPDF: A Deep Learning Approach to Analyzing PDFs. 2018 Workshop on Mining Scientific Publications at the Eleventh International Conference on Language Resources and Evaluation (LREC 2018).

INVITED FEATURE SECTION

Hepatocellular nodules in liver cirrhosis: hemodynamic evaluation (angiography-assisted CT) with special reference to multi-step hepatocarcinogenesis

Osamu Matsui, Satoshi Kobayashi, Junichiro Sanada, Wataru Kouda, Yasuji Ryu, Kazuto Kozaka, Azusa Kitao, Koichi Nakamura, Toshifumi Gabata

Department of Radiology, Graduate School of Medical Science, Kanazawa University, Kanazawa, Japan

Abstract

To understand the hemodynamics of hepatocellular carcinoma (HCC) is important for the precise imaging diagnosis and treatment, because there is an intense correlation between their hemodynamics and pathophysiology. Angiogenesis such as sinusoidal capillarization and unpaired arteries shows gradual increase during multi-step hepatocarcinogenesis from high-grade dysplastic nodule to classic hypervascular HCC. In accordance with this angiogenesis, the intranodular portal supply is decreased, whereas the intranodular arterial supply is first decreased during the early stage of hepatocarcinogenesis and then increased in parallel with increasing grade of malignancy of the nodules. On the other hand, the main drainage vessels of hepatocellular nodules change from hepatic veins to hepatic sinusoids and then to portal veins during multi-step hepatocarcinogenesis, mainly due to disappearance of the hepatic veins from the nodules. Therefore, in early HCC, no perinodular corona enhancement is seen on portal to equilibrium phase CT, but it is definite in hypervascular classical HCC. Corona enhancement is thicker in encapsulated HCC and thin in HCC without pseudocapsule. To understand these hemodynamic changes during multi-step hepatocarcinogenesis is important, especially for early diagnosis and treatment of HCCs.

Key words: Hepatocellular carcinoma—Blood supply—Multi-step hepatocarcinogenesis—Early HCC—Dysplastic nodule—Liver

Hepatocellular carcinoma (HCC) is the most common primary liver cancer worldwide. Approximately 80% of Japanese HCC cases are derived from HCV-associated liver cirrhosis and chronic hepatitis, and the remaining less than 20% of the patients are HBV positive. The patients with hepatitis B or C cirrhosis are especially classified as a very high-risk group. Ultrasonography is performed every 3–4 months for the very high-risk group. Because of the introduction of this surveillance system, the size of HCCs firstly detected during 2002–2003 ($n = 33731$) was less than 2 cm in 32.5% of all cases, 2.1–5.0 cm 47.0%, respectively [1]. However, various types of hepatocellular nodules such as dysplastic nodule (DN) are also detected during screening procedures. Ultrasound and CT features of DN and early HCCs are similar, and a precise differential diagnosis is impossible. Pathologically, human HCC develops in a multistep fashion from DN to classic hypervascular HCC. Therefore, for the early diagnosis of HCC, understanding of the concept of multi-step hepatocarcinogenesis and the sequential changes of imaging findings in accordance with multi-step hepatocarcinogenesis is important.

To understand the hemodynamics of HCC is important for the precise imaging diagnosis and treatment, because there is an intense correlation between its hemodynamic and pathophysiology. For this purpose, dynamic MDCT is most valuable because of its high

DeepPDF object detection

Table detection

2 Pathology Research International

The major pitfall of this procedure is that fine needle aspiration cytology cannot differentiate between follicular adenoma and follicular carcinoma [5, 7].

Histopathology of the excised specimen showed multinodular goiter as the commonest lesion. In one study, fine needle aspiration cytology and thyroid scan offers the best preoperative assessment of solitary thyroid nodule. Histopathology later on confirms the preoperative diagnosis. As the incidence of solitary thyroid nodule is high in Pakistan, this study will help in early detection of thyroid lesions [8].

Many modalities are used for the diagnosis of solitary nodule in Pakistan but histopathology remained the gold standard for the comparison of all. In a recent study it has been seen that FNAC is a primary investigation of thyroid lesions and used in a patient with one or more thyroid nodules. FNAC is also advised for every patient for exclusion of cancer and in the initial management of patients. It is frequently used because it is inexpensive, sensitive, specific, and an accurate procedure; therefore it is adapted as an initial investigation of thyroid diseases in all tertiary hospitals in developing countries like Pakistan [9]. In Pakistan, before performing manual or guided FNAC, thyroid scan and thyroid function tests (TFT) are performed for further management of solitary thyroid nodules. In this study we compared the individual efficacy of FNAC and thyroid scan in the management of solitary thyroid nodules. The nodules were marked after doing their scan by the nuclear center of Mayo Hospital taking histopathology as gold standard.

2. Material and Methods

This was a comparative cross-sectional study and carried out at the Departments of Surgery, Nuclear Medicine, and Pathology, King Edward Medical University, Lahore. It was conducted on 50 patients of solitary thyroid nodule in 06 months. A nonprobability purposive sampling technique was used for these patients.

Inclusion Criteria.

- (1) Age 10 to 70 Years.
- (2) Both genders.
- (3) Patient presenting with solitary swelling arising from any lobe of thyroid selected by clinical palpation.

Exclusion Criteria.

- (1) Patients with diffuse thyroid swelling.
- (2) All toxic and multinodular goiters confirmed by clinical evaluation.
- (3) Patients with history of any type of thyroid surgery (lobectomy or total thyroidectomy).

2.1. Data Collection.

All patients presenting with solitary thyroid nodules in the OPD and fulfilling the inclusion criteria were included in this study. Informed consent from all the patients included in the study was taken. All the patients were recorded for their demographic features, that is, age, sex, and address (for followup). History of present illness with regard to symptoms and duration was recorded. They were examined for the signs related to the solitary thyroid swelling. All routine investigations and serum T3, T4, and TSH levels were performed by Radioimmunoassay (RIA), (normal range of T3, 2.5–5.8 pmol/L, T4 11.5–23.0 pmol/L, and TSH, 0.2–4.0 mIU/L). Patients with solitary thyroid swelling underwent thyroid scan. Solitary thyroid nodules were selected and then FNAC was performed. The radiocollid scan and FNA were performed in conjunction as compared to FNAC alone.

Cytological diagnosis was categorized into four groups: negative for malignancy, indeterminate (suspicious) for malignancy, positive for malignancy, and inadequate. The cases were operated and evaluated for histopathological changes. The results of thyroid scan, fine needle aspiration cytology, and histopathology were compared. Histopathology was taken as gold standard. All these information were collected through proforma that is attached herewith.

2.2. Statistical Analysis.

All the data was analyzed with SPSS version 11. The variables included were demographic information, routine investigations, thyroid scan, and thyroid function tests. For quantitative data, that is, thyroid function tests, duration and size of thyroid nodule, mean, and standard deviation were calculated. For qualitative data, that is, results of thyroid scan, fine needle aspiration cytology, and histopathology, percentages was calculated. A 2 × 2 table was used to calculate sensitivity, specificity, positive predictive value, negative predictive value, and accuracy.

3. Results

The age of patients ranged from 10 to 70 years with mean age 33.04 ± 12.29 years. 41 patients (82%) were females, and 9 (18%) were males (male to female ratio 1:4.6) (Table 1).

Regarding thyroid function tests, 48 patients (96%) were euthyroid, 2 patients (4%) were hyperthyroid, and no patient was hypothyroid. The mean for serum T3, serum T4, and serum TSH were 4.00 ± 0.90, 16.47 ± 3.18, and 0.95 ± 0.85, respectively, (normal values given in Table 2). There were 40 patients (80%) who had cold nodule on thyroid scan of which 8 patients (16%) were male and 32 patients (64%)

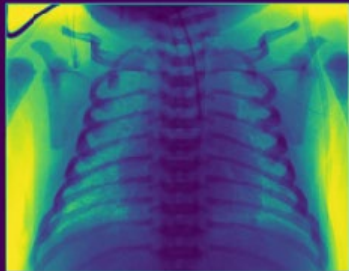
Age	Number	Percentage
Male: Female = 9:41 (18% and 82%)		
10–20	04	8.0
21–30	22	44.0
31–40	14	28.0
41–50	04	8.0
51–60	05	10.0
61–70	01	2.0
Total	50	100.0
Mean ± SD	33.04 ± 12.29	
Male to female ratio	1:4.5	

virtually complete PV obstruction occurs in the uterus. However, this anomaly should be clearly distinguished from the stenosis of aberrant PV at their junction with the systemic circulation common in TAPVD cases after birth or surgery. Prenatally, the small pulmonary venous blood flow can be diverted to the systemic veins, for example the azygos and hemiazygos veins, by way of bronchopulmonary anastomoses in the lungs,^{1,16–19} and persistently dilated fetal pulmonary lymphatic channels help to remove the excess interstitial edema fluid from the lungs.^{2,7,23,24} Therefore, the fetus possibly tolerates severe obstruction to pulmonary venous drainage. In contrast, postnatally, with the tenfold increase in pulmonary blood flow, these channels are very easily overwhelmed.⁸ Neonates with ACPV generally display symptoms of cardiorespiratory failure, such as cyanosis and/or acidosis within minutes of birth.^{1,25–31} Pulmonary interstitial edema or occasional secondary pulmonary lymphangiectasis (PL) also impairs the gaseous exchange.^{25,32,33} Cardiovascular and respiratory collapse and subsequent death rapidly occur in many ACPV cases, and early recognition and surgical intervention are essential to patient care.^{1,25–31} According to previous reports, however, only five cases have undergone successful surgical repair.^{2,7}

In this paper we describe an autopsy case of TAPVD type I, in which anomalous small right PV drained to the superior vena cava (SVC), associated with ACPV of the left and right inferior PV (IPV) with blind confluence and atresia of the left superior PV (SPV). In addition, secondary PL was also identified in the histological examination.

CLINICAL SUMMARY

A Japanese male neonate was born at 37 weeks of gestation without history of pregestational treatment for infertility through Caesarean section due to poor fetal movement for 2 days before his birth. He presented with apnea immediately after birth and his Apgar scores were very low, 1 and 2 points at 1 and 5 min, respectively. He was the second child born to non-consanguineous parents who were both 40 years old. The parents' first baby was also born at full term of spontaneous gestation 2 years before this case, and had weight appropriate for gestational age. This patient, however, weighed about 2420 g at birth. The amniotic fluid was clear, and the placenta, weighing 420 g, had no remarkable changes. A chest X-ray of the case showed marked pulmonary congestion and diffuse bilateral reticulogranular infiltrates in the lungs without evidence of cardiomegaly (Fig. 1) and an echocardiogram revealed that PV did not connect to the left atrium or CPV but an anomalous vein-like small vein drained into the SVC, suggestive of TAPVD, supracardiac type. The patient could not have surgery because of his poor general condition. Although artificial ventilation after intuba-



tion and surfactant substitution therapy were performed, he died of hypoxemic respiratory failure, which was resistant to the intravenous administration of dopamine and bicarbonate and the inhalation of nitric oxide, 6 days after birth.

The blood cell counts and values of blood biochemistry showed thrombocytopenia ($6.8 \times 10^9/\mu\text{L}$), high lactate acidemia (108 mg/dL) and severe hypoglycemia (4 mg/dL). Additionally, a venous blood sample when breathing 100% oxygen concentration in the incubator demonstrated elevated partial pressure of carbon dioxide (PaCO_2 : 73.6 mmHg), abd decreased base excess (-10.2 mEq/L) and pH (7.097), indicating marked respiratory acidosis. An autopsy was performed approximately 20 h after death, but the brain could not be examined due to the family's objections.

PATHOLOGICAL FINDINGS

At autopsy, the baby measured approximately 46 cm in height and weighed 2420 g. An external examination showed anomaly of the right accessory aricle and a few needle marks on the extremities. In an internal examination of the thoracic cavity, markedly involuted thymus, weighing 1.5 g, was noted in the anterior mediastinum. The heart, weighing 21.5 g, showed a dilated and hypertrophic right ventricle and modest hypoplasia of the left ventricle without apparent PV returning to the left atrium (Fig. 2a), consistent with the findings of TAPVD. Although there were no PV draining into the right atrium or lunoculi, reminiscent of an atretic fibrous band of CPV, between the atrium and pericardium (Fig. 2g) two IPV from the lower lobes of the lungs were seen to form a

94 S. Yamada et al.

Paragraph detection

virtually complete PV obstruction occurs in the uterus. However, this anomaly should be clearly distinguished from the stenosis of aberrant PV at their junction with the systemic circulation common in TAPVD cases after birth or surgery. Prenatally, the small pulmonary venous blood flow can be diverted to the systemic veins, for example the azygos and hemiazygos veins, by way of bronchopulmonary anastomoses in the lungs,^{1,16–19} and persistently dilated fetal pulmonary lymphatic channels help to remove the excess interstitial edema fluid from the lungs.^{2,7,23,24} Therefore, the fetus possibly tolerates severe obstruction to pulmonary venous drainage. In contrast, postnatally, with the tenfold increase in pulmonary blood flow, these channels are very easily overwhelmed.⁸ Neonates with ACPV generally display symptoms of cardiorespiratory failure, such as cyanosis and/or acidosis within minutes of birth.^{1,25–31} Pulmonary interstitial edema or occasional secondary pulmonary lymphangiectasis (PL) also impairs the gaseous exchange.^{25,32,33} Cardiovascular and respiratory collapse and subsequent death rapidly occur in many ACPV cases, and early recognition and surgical intervention are essential to patient care.^{1,25–31} According to previous reports, however, only five cases have undergone successful surgical repair.^{2,7}

In this paper we describe an autopsy case of TAPVD type I, in which anomalous small right PV drained to the superior vena cava (SVC), associated with ACPV of the left and right inferior PV (IPV) with blind confluence and atresia of the left superior PV (SPV). In addition, secondary PL was also identified in the histological examination.

CLINICAL SUMMARY

A Japanese male neonate was born at 37 weeks of gestation without history of pregestational treatment for infertility through Caesarean section due to poor fetal movement for 2 days before his birth. He presented with apnea immediately after birth and his Apgar scores were very low, 1 and 2 points at 1 and 5 min, respectively. He was the second child born to non-consanguineous parents who were both 40 years old. The parents' first baby was also born at full term of spontaneous gestation 2 years before this case, and had weight appropriate for gestational age. This patient, however, weighed about 2420 g at birth. The amniotic fluid was clear, and the placenta, weighing 420 g, had no remarkable changes. A chest X-ray of the case showed marked pulmonary congestion and diffuse bilateral reticulogranular infiltrates in the lungs without evidence of cardiomegaly (Fig. 1) and an echocardiogram revealed that PV did not connect to the left atrium or CPV but an anomalous vein-like small vein drained into the SVC, suggestive of TAPVD, supracardiac type. The patient could not have surgery because of his poor general condition. Although artificial ventilation after intuba-



tion and surfactant substitution therapy were performed, he died of hypoxemic respiratory failure, which was resistant to the intravenous administration of dopamine and bicarbonate and the inhalation of nitric oxide, 6 days after birth.

The blood cell counts and values of blood biochemistry showed thrombocytopenia ($6.8 \times 10^9/\mu\text{L}$), high lactate acidemia (108 mg/dL) and severe hypoglycemia (4 mg/dL). Additionally, a venous blood sample when breathing 100% oxygen concentration in the incubator demonstrated elevated partial pressure of carbon dioxide (PaCO_2 : 73.6 mmHg), abd decreased base excess (-10.2 mEq/L) and pH (7.097), indicating marked respiratory acidosis. An autopsy was performed approximately 20 h after death, but the brain could not be examined due to the family's objections.

PATHOLOGICAL FINDINGS

At autopsy, the baby measured approximately 46 cm in height and weighed 2420 g. An external examination showed anomaly of the right accessory aricle and a few needle marks on the extremities. In an internal examination of the thoracic cavity, markedly involuted thymus, weighing 1.5 g, was noted in the anterior mediastinum. The heart, weighing 21.5 g, showed a dilated and hypertrophic right ventricle and modest hypoplasia of the left ventricle without apparent PV returning to the left atrium (Fig. 2a), consistent with the findings of TAPVD. Although there were no PV draining into the right atrium or lunoculi, reminiscent of an atretic fibrous band of CPV, between the atrium and pericardium (Fig. 2g) two IPV from the lower lobes of the lungs were seen to form a

© 2010 The Authors
Pathology International © 2010 Japanese Society of Pathology and Blackwell Publishing Ltd

Next steps – table structure prediction

- Identify rows, columns, headers, and other structure
- Example below: two successful (a, b) and two challenging (c, d) cases

(a) Row detection, no ruling lines present

(b) Column detection, no ruling lines present

Sample Group	Some Year 1 Head Start Participation	No Year 1 Head Start Participation	Total
All Randomly Assigned (N=4,667):			
3-Year-Old Cohort			
Head Start Group	85.1%	14.9%	100%
Control Group	17.3%	82.7%	100%
4-Year-Old Cohort			
Head Start Group	79.8%	20.2%	100%
Control Group	13.9%	86.1%	100%

(c) Nested row hierarchy

Table A-1. Summary of forecast assumptions to 2021

Variable	Assumption
Demographic assumptions	Projections are consistent with the Census Bureau estimates
Population	Census Bureau projection: average annual growth rate of 0.1%
15- to 24-year-old population	Census Bureau projection: average annual growth rate of 0.6%
25- to 34-year-old population	Census Bureau projection: average annual growth rate of 1.3%
35- to 44-year-old population	Census Bureau projection: average annual growth rate of 0.6%
Economic assumptions	Annual percent changes range between -1.3% and 2.2%
Disposable income per capita in constant dollars	with an annual growth rate of 1.4%
Education revenue receipts from state sources per capita in constant dollars	Annual percent changes range between -2.4% and 2.3%
Inflation rate	with an annual growth rate of 1.2%
	Inflation rate ranges between 1.0% and 2.0%
Unemployment rate (men)	Remains between 17.7% and 26.8%
Ages 16 and 19	Remains between 10.8% and 15.6%
Ages 20 to 24	Remains between 5.3% and 7.9%
Age 25 and over	
Unemployment rate (women)	Remains between 14.3% and 19.6%
Ages 16 and 19	Remains between 9.3% and 13.1%
Ages 20 to 24	Remains between 5.0% and 7.3%
Age 25 and over	

* As the Census Bureau projections were not updated to reflect the 2011 Census Bureau population estimates, the Census Bureau age-specific

(d) Very close-by rows

Source: Schreiber et al. (2017). DeepDeSRT: Deep Learning for Detection and Structure Recognition of Tables in Document Images. ICDAR 2017.

Next steps – text extraction

- Correctly extract text contained in tables
- Older documents particularly challenging
- Example extraction using Tabula (<https://tabula.technology/>)

Table 1. Uterine and vaginal responses to isoflavones and diethylstilboestrol, and relative potencies of biochanin A, genistein and diethylstilboestrol

Material	Total dose /mouse (six mice/group)	Mean uterine weight (mg.)	No. of positive vaginal smears
Controls	—	5.9 ± 0.13*	0
Formononetin	3.00 mg.	6.1 ± 0.13	0
	5.99 mg.	6.4 ± 0.13	0
	11.96 mg.	6.7 ± 0.13	0
Genistein	2.99 mg.	6.8	0
	5.77 mg.	12.3	5
	11.13 mg.	23.3	5
Biochanin A	2.94 mg.	7.0	0
	5.90 mg.	8.2	4
	11.72 mg.	14.2	5
Diethylstilboestrol	0.036 µg.	7.9	0
	0.072 µg.	9.7	0
	0.144 µg.	16.0	0

0.036 µg.
0.072 µg.
0.144 µg.

Relative activities based on uterine weight response (95% fiducial limits): genistein/biochanin A, 1.49 (0.97–2.29); genistein/diethylstilboestrol, 13.5×10^{-6} (10.3 – 17.8×10^{-6}); biochanin A/diethylstilboestrol, 9.1×10^{-6} (6.0 – 13.7×10^{-6}).

* s.e.m.'s apply to control and formononetin treated mice only.

Material	Total dose /mouse (six mice/group)	Mean uterine weight (mg.)	No. of positive vaginal smears
Controls	—	5.9 ± 0.13*	0
Formononetin	3.00 mg.	6.1 ± 0.13	0
	5.99 mg.	6.4 ± 0.13	0
	11.96 mg.	6.7 ± 0.13	0
Genistein	2.99 mg.	6.8	0
	5.77 mg.	12.3	5
	11.13 mg.	23.3	5
Biochanin A	2.94 mg.	7.0	0
	5.90 mg.	8.2	4
	11.72 mg.	14.2	5
Diethylstilboestrol	0.036 pg.	7.9	0
	0.072 fig.	9.7	0
	0.144 pg.	16.0	0

0.036 pg.

0.072 fig.

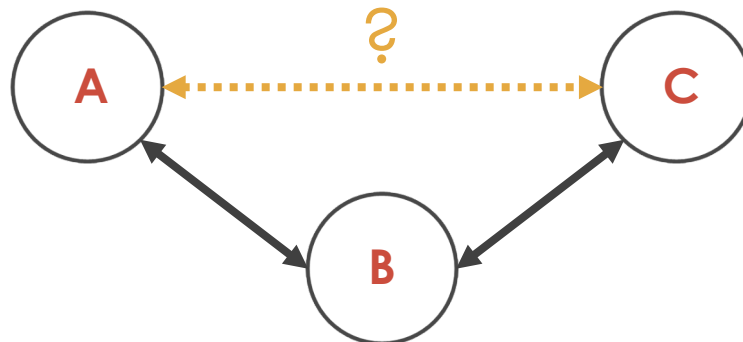
0.144 pg.

Towards deriving actionable insights from scientific knowledge graphs



Research challenge: Literature-based discovery (LBD)

- LBD is the process of **knowledge extraction** and **automated hypothesis generation** from scientific literature
- Traditional approach: Swanson Linking
 - There are published relations between A and B, and B and C
 - **There may be a yet undiscovered relation between A-C**
- Demonstrated on the connection between dietary fish oil (A) and Raynaud's disease (C) through high blood viscosity (B) [1]

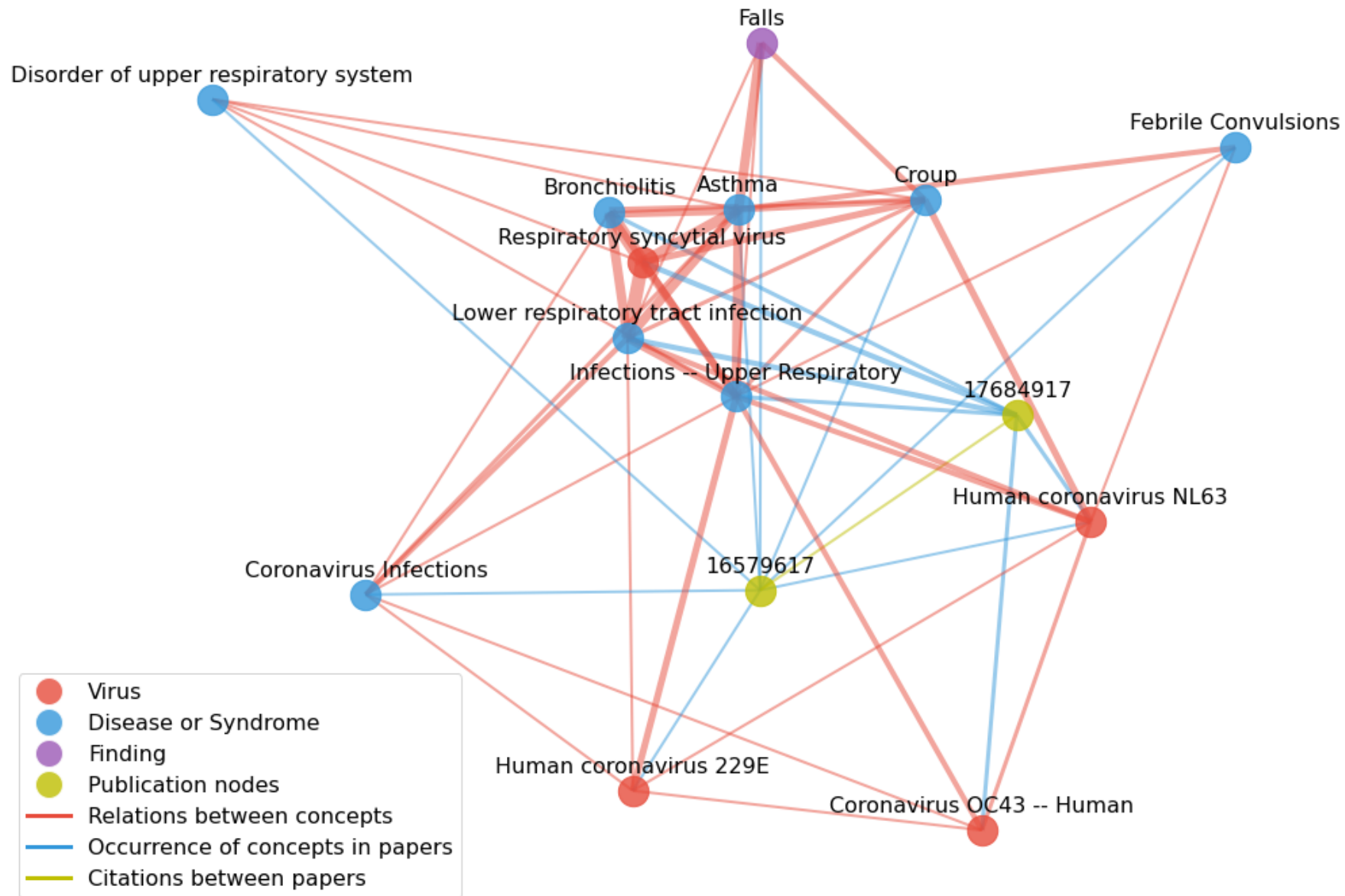


Approach

- Previous works hypothesized **shortest path information may be used for automating literature-based discovery**
- To test this hypothesis we created a **graph of A->B->C connections** from Semantic MEDLINE + CORON-19 publications
 - **Graph nodes (18.5 million total)**
 - Biomedical concepts (UMLS)
 - Publications
 - **Graph relations (213 million total):**
 - Relations between **concepts** (extracted from publication abstracts)
 - Relations between **concepts** and **publications** (concept appears in paper)
 - Citations between **publications**

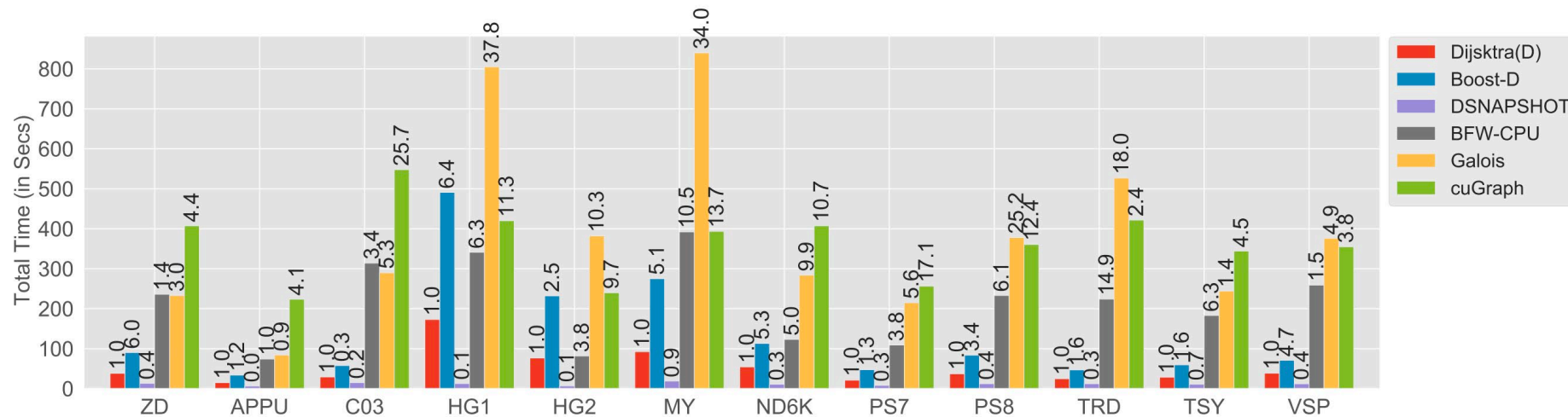


Graph example



All Pairs Shortest Path algorithm – DSNAPSHOT

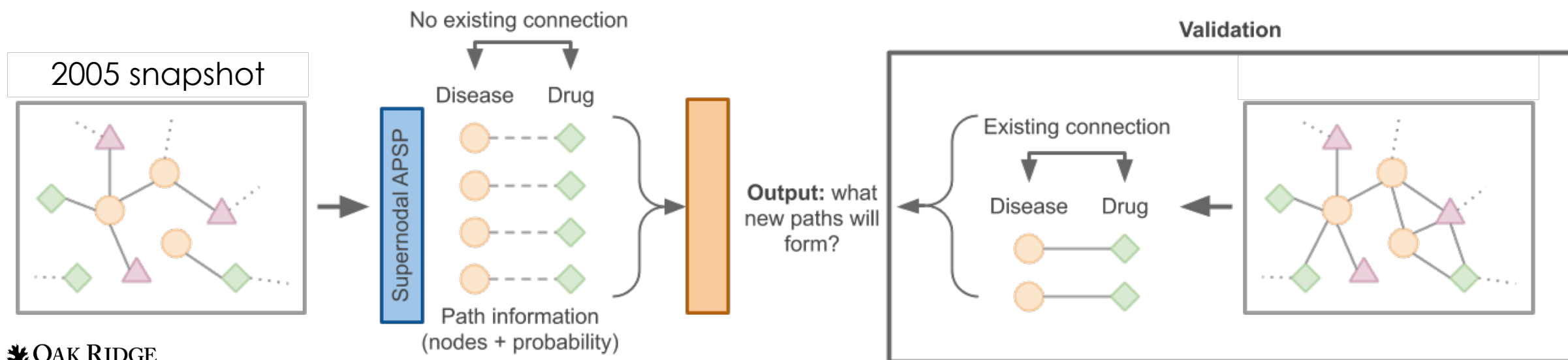
- **Goal:** Find the shortest paths between all nodes within an 18 million node graph
- **Approach:** Leverage graph sparsity and algorithmic optimizations to enable GPU usage on ORNL Summit [2, 3]
 - 18 hrs to process entire graph – **10X speed up over the state of the art**



2. Sao P, Kannan R, Gera P and Vuduc R. A supernodal all-pairs shortest path algorithm. In Proceedings of the 25th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming. 2020. 250-261.
3. R. Kannan et al. Scalable Knowledge-Graph Analytics at 136 Petaflops/s. To appear in Proceedings of the ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis (IEEE Supercomputing). 2020.

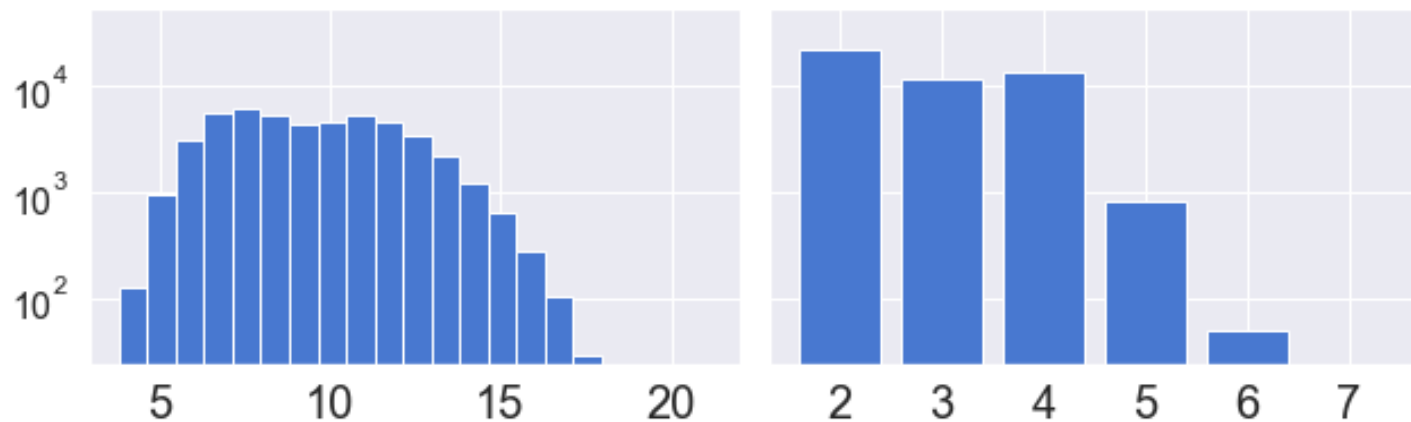
DSNAPSHOT 2005 Experiment

- Use **DSNAPSHOT** on PubMed data from 1970 – 2005 (Done)
- Use **DSNAPSHOT results** to recommend new connections (A->C) for 2005 (Done)
- Validate these connections using 2020 data (in progress)
- Create a ranking method for new potential connections (in progress)

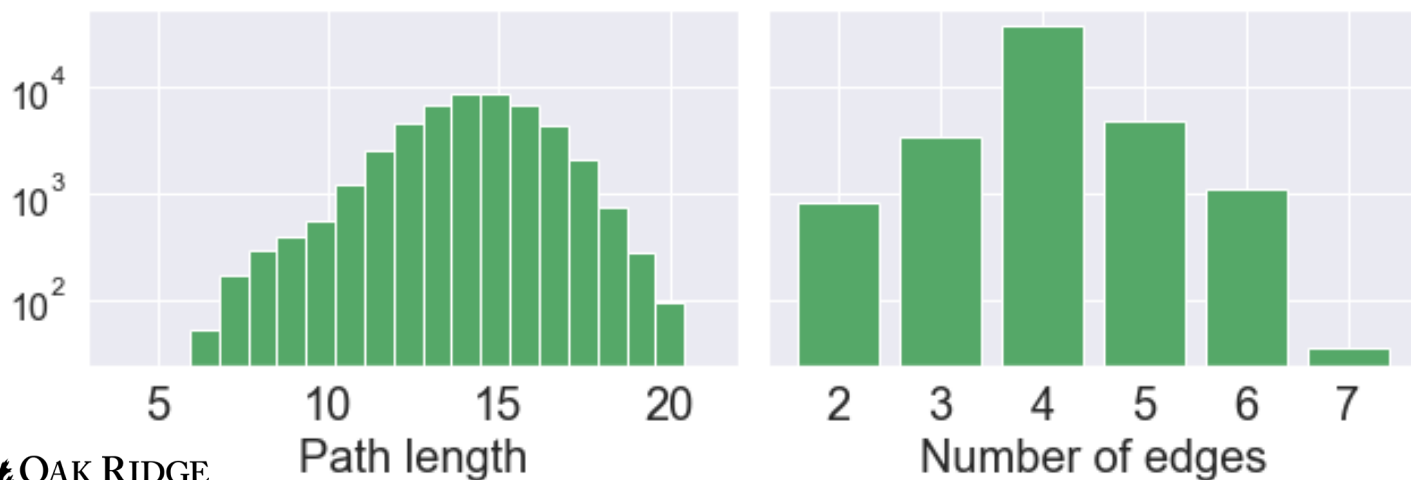


Results of 2005 - Learning to Discover

Connections discovered by 2020



No connections by 2020



- Paths that lead to a later connection are different from paths that don't lead to a connection
- Hypothesis: This is a ranking method for predicting future discoveries

Example results in **2005** snapshot

Successful outcome of human metapneumovirus (hMPV) pneumonia in a lung transplant recipient treated with intravenous ribavirin (2008):

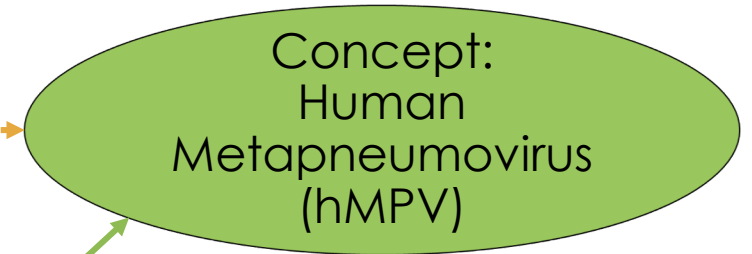
hMPV was diagnosed by polymerase chain reaction, and treated with intravenous **ribavirin** with a successful outcome.

Pharmacologic Substance

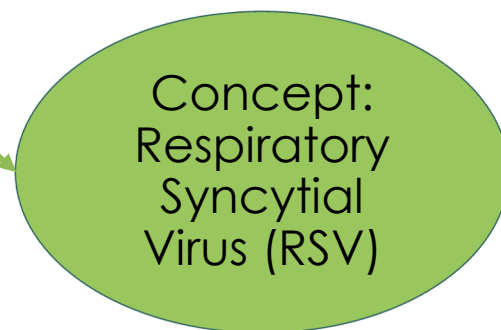


Is there a connection?

Virus



Virus



Ribavirin and Respiratory Syncytial Virus (1986):

Ribavirin [...] possesses an unusually broad spectrum of antiviral activity, inhibiting under laboratory conditions a wide variety of RNA and DNA viruses, including among respiratory viruses influenza types A and B, parainfluenza types 1, 2, and 3, **respiratory syncytial virus (RSV)**, and possibly coronavirus.

The causes and diagnosis of influenza-like illness (2004):

In addition to influenza, viruses known to cause [influenza-like illness] include **respiratory syncytial virus**, rhinovirus, adenovirus, parainfluenza viruses, human coronaviruses [...] and the recently recognised **human metapneumovirus**.

Acknowledgements and contact



Acknowledgements

- NIEHS/NTP
 - Mary Wolfe
 - Nicole Kleinstreuer
 - Charles Schmitt
- Oak Ridge National Laboratory
 - Steven Young
 - Christopher Stahl
 - Chelsey Stahl
 - Dasha Herrmannova



Acknowledgements

- Support for this research was provided by an Interagency Agreement with the National Institute of Environmental Health Sciences (AES 16002-001) and the U.S. Department of Energy at Oak Ridge National Laboratory.
- This manuscript has been authored by UT-Battelle, LLC and used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan.

Contact

Robert M. Patton, Ph.D.

Computer Science & Mathematics Division

Oak Ridge National Laboratory

pattonrm@ornl.gov

Extra Content



Results for MC 1: Animal model

The intact female weanling version in the Organization for Economic Cooperation and Development (OECD) uterotrophic assay Test Guideline (TG) 440 is proposed as an alternative to the adult ovariectomized female version, because it does not involve surgical intervention (vs the ovariectomized version) and detects direct/indirect-acting estrogenic/anti-estrogenic substances (vs the ovariectomized version which detects only direct-acting estrogenic/anti-estrogenic substances binding to the estrogen receptor). This validation study followed OECD TG 440, with six female weanling rats (postnatal day 21) per dose group and six treatment groups. Females were weighed and dosed once daily by oral gavage for three consecutive days, with one of six doses of 17alpha-ethinyl estradiol in corn oil at 5 ml kg⁻¹ at 0 and 0.1-10 microg kg⁻¹ per day. On postnatal day 24, the juvenile females were euthanized by CO₂ asphyxiation, weighed, livers weighed and uteri weighed wet and blotted. The presence or absence of vaginal patency was recorded. Absolute and relative (to terminal body weight) uterine wet and blotted weights and uterine luminal fluid weights were significantly increased at 3.0 and 10.0 (both P < 0.01) microg kg⁻¹ per day, and increased to ~140% of control values at 1.0 microg kg⁻¹ per day (not statistically significantly). In vivo body weights, weight changes, feed consumption, liver weights and terminal body weights were unaffected. Vaginal patency was not acquired in any female at any dose, although vaginal puckering was observed in one female at 10.0 microg kg⁻¹ per day. Therefore, this intact weanling uterotrophic assay is validated in our laboratory for use under US and European endocrine toxicity testing programs/legislation.

← Correct answer
(underlined)

Results for MC 2: Group size

The intact female weanling version in the Organization for Economic Cooperation and Development (OECD) uterotrophic assay Test Guideline (TG) 440 is proposed as an alternative to the adult ovariectomized female version, because it does not involve surgical intervention (vs the ovariectomized version) and detects direct/indirect-acting estrogenic/anti-estrogenic substances (vs the ovariectomized version which detects only direct-acting estrogenic/anti-estrogenic substances binding to the estrogen receptor). This validation study followed OECD TG 440, with six female weanling rats (postnatal day 21) per dose group and six treatment groups. Females were weighed and dosed once daily by oral gavage for three consecutive days, with one of six doses of 17alpha-ethinyl estradiol in corn oil at 5 ml kg⁻¹ at 0 and 0.1-10 microg kg⁻¹ per day. On postnatal day 24, the juvenile females were euthanized by CO₂ asphyxiation, weighed, livers weighed and uteri weighed wet and blotted. The presence or absence of vaginal patency was recorded. Absolute and relative (to terminal body weight) uterine wet and blotted weights and uterine luminal fluid weights were significantly increased at 3.0 and 10.0 (both P < 0.01) microg kg⁻¹ per day, and increased to ~140% of control values at 1.0 microg kg⁻¹ per day (not statistically significantly). In vivo body weights, weight changes, feed consumption, liver weights and terminal body weights were unaffected. Vaginal patency was not acquired in any female at any dose, although vaginal puckering was observed in one female at 10.0 microg kg⁻¹ per day. Therefore, this intact weanling uterotrophic assay is validated in our laboratory for use under US and European endocrine toxicity testing programs/legislation.

←
Correct answer
(underlined)

Results for MC 3: Route of administration

The intact female weanling version in the Organization for Economic Cooperation and Development (OECD) uterotrophic assay Test Guideline (TG) 440 is proposed as an alternative to the adult ovariectomized female version, because it does not involve surgical intervention (vs the ovariectomized version) and detects direct/indirect-acting estrogenic/anti-estrogenic substances (vs the ovariectomized version which detects only direct-acting estrogenic/anti-estrogenic substances binding to the estrogen receptor). This validation study followed OECD TG 440, with six female weanling rats (postnatal day 21) per dose group and six treatment groups. Females were weighed and dosed once daily by oral gavage for three consecutive days, with one of six doses of 17alpha-ethinyl estradiol in corn oil at 5 ml kg⁻¹ at 0 and 0.1-10 microg kg⁻¹ per day. On postnatal day 24, the juvenile females were euthanized by CO₂ asphyxiation, weighed, livers weighed and uteri weighed wet and blotted. The presence or absence of vaginal patency was recorded. Absolute and relative (to terminal body weight) uterine wet and blotted weights and uterine luminal fluid weights were significantly increased at 3.0 and 10.0 (both P < 0.01) microg kg⁻¹ per day, and increased to ~140% of control values at 1.0 microg kg⁻¹ per day (not statistically significantly). In vivo body weights, weight changes, feed consumption, liver weights and terminal body weights were unaffected. Vaginal patency was not acquired in any female at any dose, although vaginal puckering was observed in one female at 10.0 microg kg⁻¹ per day. Therefore, this intact weanling uterotrophic assay is validated in our laboratory for use under US and European endocrine toxicity testing programs/legislation.



Correct answer
(underlined)

Results for MC 4: Number of dose groups

The intact female weanling version in the Organization for Economic Cooperation and Development (OECD) uterotrophic assay Test Guideline (TG) 440 is proposed as an alternative to the adult ovariectomized female version, because it does not involve surgical intervention (vs the ovariectomized version) and detects direct/indirect-acting estrogenic/anti-estrogenic substances (vs the ovariectomized version which detects only direct-acting estrogenic/anti-estrogenic substances binding to the estrogen receptor). This validation study followed OECD TG 440, with six female weanling rats (postnatal day 21) per dose group and six treatment groups. Females were weighed and dosed once daily by oral gavage for three consecutive days, with one of six doses of 17alpha-ethinyl estradiol in corn oil at 5 ml kg⁻¹ at 0 and 0.1-10 microg kg⁻¹ per day. On postnatal day 24, the juvenile females were euthanized by CO₂ asphyxiation, weighed, livers weighed and uteri weighed wet and blotted. The presence or absence of vaginal patency was recorded. Absolute and relative (to terminal body weight) uterine wet and blotted weights and uterine luminal fluid weights were significantly increased at 3.0 and 10.0 (both P < 0.01) microg kg⁻¹ per day, and increased to ~140% of control values at 1.0 microg kg⁻¹ per day (not statistically significantly). In vivo body weights, weight changes, feed consumption, liver weights and terminal body weights were unaffected. Vaginal patency was not acquired in any female at any dose, although vaginal puckering was observed in one female at 10.0 microg kg⁻¹ per day. Therefore, this intact weanling uterotrophic assay is validated in our laboratory for use under US and European endocrine toxicity testing programs/legislation.



Correct answer
(underlined)

Results for MC 5: Dosing interval

The intact female weanling version in the Organization for Economic Cooperation and Development (OECD) uterotrophic assay Test Guideline (TG) 440 is proposed as an alternative to the adult ovariectomized female version, because it does not involve surgical intervention (vs the ovariectomized version) and detects direct/indirect-acting estrogenic/anti-estrogenic substances (vs the ovariectomized version which detects only direct-acting estrogenic/anti-estrogenic substances binding to the estrogen receptor). This validation study followed OECD TG 440, with six female weanling rats (postnatal day 21) per dose group and six treatment groups. Females were weighed and dosed once daily by oral gavage for three consecutive days, with one of six doses of 17alpha-ethinyl estradiol in corn oil at 5 ml kg⁻¹ at 0 and 0.1-10 microg kg⁻¹ per day. On postnatal day 24, the juvenile females were euthanized by CO₂ asphyxiation, weighed, livers weighed and uteri weighed wet and blotted. The presence or absence of vaginal patency was recorded. Absolute and relative (to terminal body weight) uterine wet and blotted weights and uterine luminal fluid weights were significantly increased at 3.0 and 10.0 (both P < 0.01) microg kg⁻¹ per day, and increased to ~140% of control values at 1.0 microg kg⁻¹ per day (not statistically significantly). In vivo body weights, weight changes, feed consumption, liver weights and terminal body weights were unaffected. Vaginal patency was not acquired in any female at any dose, although vaginal puckering was observed in one female at 10.0 microg kg⁻¹ per day. Therefore, this intact weanling uterotrophic assay is validated in our laboratory for use under US and European endocrine toxicity testing programs/legislation.



Correct answer
(underlined)

Results for MC 6: Necropsy timing

The intact female weanling version in the Organization for Economic Cooperation and Development (OECD) uterotrophic assay Test Guideline (TG) 440 is proposed as an alternative to the adult ovariectomized female version, because it does not involve surgical intervention (vs the ovariectomized version) and detects direct/indirect-acting estrogenic/anti-estrogenic substances (vs the ovariectomized version which detects only direct-acting estrogenic/anti-estrogenic substances binding to the estrogen receptor). This validation study followed OECD TG 440, with six female weanling rats (postnatal day 21) per dose group and six treatment groups. Females were weighed and dosed once daily by oral gavage for three consecutive days, with one of six doses of 17alpha-ethinyl estradiol in corn oil at 5 ml kg⁽⁻⁾(1) at 0 and 0.1-10 microg kg⁽⁻⁾(1) per day. On postnatal day 24, the juvenile females were euthanized by CO₂ asphyxiation, weighed, livers weighed and uteri weighed wet and blotted. The presence or absence of vaginal patency was recorded. Absolute and relative (to terminal body weight) uterine wet and blotted weights and uterine luminal fluid weights were significantly increased at 3.0 and 10.0 (both P < 0.01) microg kg⁽⁻⁾(1) per day, and increased to ~140% of control values at 1.0 microg kg⁽⁻⁾(1) per day (not statistically significantly). In vivo body weights, weight changes, feed consumption, liver weights and terminal body weights were unaffected. Vaginal patency was not acquired in any female at any dose, although vaginal puckering was observed in one female at 10.0 microg kg⁽⁻⁾(1) per day. Therefore, this intact weanling uterotrophic assay is validated in our laboratory for use under US and European endocrine toxicity testing programs/legislation.



Correct answer
(underlined)