ISSN 2682- 6925 (online)

JOURNAL OF JAFFNA SCIENCE ASSOCIATION

Volume 1 Issue 1

JAFFNA SCIENCE ASSOCIATION, SRI LANKA
2019

JOURNAL OF JAFFNA SCIENCE ASSOCIATION

Volume 1 Issue 1

April 2019

JAFFNA SCIENCE ASSOCIATION, SRI LANKA

Peer Reviewed Open Access e-Journal

Journal of Jaffna Science Association (JJSA), Vol. 1, Issue 1, April 2019

https://www.journal.thejsa.org

ISSN 2682- 6925 (online)

Editorial Board of JJSA

Cover Designed By Mr.K.Harichandra

Published by:

Jaffna Science Association

No. 84, College Road, Neeraviyady, Jaffna, Sri Lanka.

e-mail: editorjjsa@thejsa.org

Webpage : https://www.thejsa.org

ISSN 2682- 6925 (online)

JOURNAL OF JAFFNA SCIENCE ASSOCIATION

Volume 1 Issue 1 April 2019

CONTENTS

Editor's Message

It gives me immense pleasure and satisfaction to release the first ever issue of the *Journal of Jaffna Science Association* (*JJSA*). The JJSA provides flat form for researchers to disseminate their research findings, mainly under local settings, in different fields of science to wider scientific community. I have the privilege to present five articles on different aspects of science to readers.

I would like to thank the Editorial Board Members and Reviewers for their valuable support and the assistance of the Jaffna Science Association in publishing this issue. A special word of appreciation goes to Mrs. R. Mahesan, Dr.(Mrs.) K. Chandarasekar and Dr. E.Y. A Chales for their untiring effort to make this issue open access. We are making every effort to accomplish the scope of JJSA.

We expect that you will continue to work with us in the future

S.N. Surendran
Professor in Zoology
University of Jaffna
Jaffna 40000

Leading Article

Folic Acid and Reproductive Health

[1]Vikram Talaulikar*, [2]Sabaratnam Arulkumaran

[1]*Reproductive Medicine Unit, University College London Hospital, United Kingdom*
[2]*Department of Obstetrics and Gynaecology, St. George's University of London, United Kingdom*

Abstract:

Folic acid is one of the B complex vitamins and is now recognised as a vital component of the peri-conceptional care of women in the reproductive age group. Deficiency of folic acid can lead to neural tube defects in the fetus and megaloblastic anaemia in the mother. Due to its lower bioavailability from natural foods, many countries have adopted mandatory folic acid food fortification programs. Although these programs have been a public health triumph in reducing the burden of neural tube defects, there have been growing concerns about the role played by folic acid supplementation in the rising colon cancer rates over the past two decades. Folate intake in the range of the dietary reference intake is associated with a reduced risk for cancer in healthy populations, whereas high intakes of folic acid might result in an increased risk for cancer incidence or progression in persons with precancerous lesions and under certain conditions. Since no adverse effects have been observed in association with the intake of dietary folate, research activities that aim at investigating cause and effect relationships focus on folic acid. The majority of the evidence available to date therefore is reassuring and until further long-term population as well as laboratory studies are completed, folic acid will continue to play an important role in pre-conceptional and early pregnancy care.

Keywords: Deficiency, Folic acid, Reproduction, Public health

1. Introduction

Over the past several decades, folic acid has been an integral part of peri-conceptional care throughout the world. Since the recognition of the link between folic acid and neural tube defects (NTDs) in 1960s, folic acid supplementation has remained the only intervention that can prevent serious congenital anomalies in the fetus.

There is now an overwhelming body of evidence that folic acid supplementation significantly reduces the risk of neural tube defects (NTDs) (De-Regil *et al*., 2010). The programs of mandatory food fortification with folic acid in various countries have been a massive public health success with regards to reduction of these anomalies (Obican *et al*., 2010). It is important to note that certain groups of patients may have an increased need for these supplements.

Recent years have also witnessed a growing controversy regarding the safety of fortification of foods with folic acid, with concerns expressed over its potential role in promoting colorectal cancer and B_{12} deficiency anaemia.

2. Folic acid – metabolism and role in human physiology

Folic acid, a water-soluble B complex vitamin derives its name from '*folium'* (Latin for leaf) as it was first extracted from spinach leaves. It was Lucy Wells in 1931 who originally used a

Corresponding author: Mr Vikram Talaulikar, MD MRCOG PhD (Associate Specialist), Reproductive Medicine Unit, University College London Hospital, 235 Euston Road, London NW1 2BU, e-mail: vtalauliker@gmail.com

yeast extract to correct macrocytic anaemia in pregnant women in India. This abstract was subsequently identified as folate (Wills, 1931). Folic acid was isolated by Mitchell *et al*. in 1941 and was shown to be a growth factor for *Streptococcus lactis R* (Mitchell *et al*., 1941). Folic acid consists of a pteridine ring system, p-aminobenzoic acid and one molecule of glutamic acid (chemical name: pteroylglutamic acid). Naturally occurring 'folates' are pteroylpolyglutamic acids with two to eight glutamic acid groups. 'Folic acid' is the synthetic version of folate and is used in vitamin supplements and fortified foods because of its increased stability. The active form of folate in the body is 'tetrahydrofolic acid', which acts as a coenzyme in numerous essential metabolic reactions. Folate coenzymes act as acceptors and donors of one-carbon units in these reactions. The synthesis of the amino acid methionine from homocysteine requires a folate coenzyme and, in addition vitamin B_{12}. Increased homocysteine levels therefore may indicate folate deficiency. Tetrahydrofolic acid is vital for synthesis as well as repair of nucleic acids (DNA and RNA) and formation of blood cells.

3. Sources of folic acid and food fortification

Natural foods like leafy green vegetables, dried beans, legumes, fruits and organ foods such as liver are rich sources of folates.

Foods that are high in folate content

- Leafy green vegetables such as spinach, kale, brussels sprouts, cabbage, broccoli, turnip greens, potatoes
- Citrus fruits such as orange, orange juice
- Beans and legumes such as peas, chickpeas, blackeye beans, lentils
- Wheat bran and other whole grain foods
- Breads
- Cereals
- Rice
- Pastas
- Poultry, pork, shellfish and liver

Most dietary folates exist as polyglutamates, which are converted to the monoglutamate form and absorbed in the proximal small intestine. However, the body absorbs only about 50 percent of food folate. This problem is compounded by cooking practices such as prolonged stewing, processing and storage, which can destroy some of the folates in natural foods.

To overcome these deficiencies, many countries have introduced food fortification of grain products including flour, rice, pasta, bread and breakfast cereals with folic acid because, in contrast to natural foods, the folic acid absorption rate is estimated at 85 percent from fortified foods and 100 percent from a vitamin supplement. In March of 1996, the United States (US) Food and Drug Administration mandated that all enriched flour and uncooked cereal grains sold in the United States should be fortified with 140 µg folic acid/100 g of flour no later than January of 1998 (Food and Drug Administration, 1996). Since then other countries including Canada, Costa Rica, Chile, Australia, New Zealand and some Middle Eastern countries have also introduced their own folic acid fortification programs. Although in 2007, the Food Standards Agency recommended that folic acid should be added compulsorily to either bread or flour to help reduce the numbers of neural tube defects, the issue of fortification still remains a subject of debate and controversy in the United Kingdom (UK). The recommended daily allowances for folates are expressed in Dietary Folate Equivalents (DFEs). DFEs were introduced because of the different bioavailability of folates and folic acid from various sources. The recommended daily allowance (RDA) for both men and women is 400 µg/day of dietary folate equivalents

(DFEs), increasing to 600 µg/day in pregnancy and 500 µg/day in lactation. DFEs adjust for the nearly 50 percent lower bioavailability of food folate compared with that of folic acid: 1 µg of dietary folate equivalent = 0.6 µg of folic acid from fortified food or as a supplement taken with meals = 0.5 µg of a supplement taken on an empty stomach. To reduce the risk of neural tube defects for women capable of becoming pregnant, the recommendation is to take 400 µg of folic acid daily from fortified foods, supplements, or both in addition to consuming food folate from a varied diet.

4. Need for folic acid supplementation

There are various contributors to folate deficiency in different populations. In pregnancy, the rising demands of increased red blood cell (RBC) production and growing fetal tissues causes a fall in blood levels of folate. The first few weeks of pregnancy are crucial for the development of the neural tube and this is the period when folic acid is most essential.

About 5—15% of the general population have a variant of 5, 10-methylenetetrahydrofolate reductase enzyme essential for catalysing the transfer of a methyl group to homocysteine to form methionine, and the presence of this variant can compromise tissue folate levels (Molloy *et al*., 1997).

Low folate intake also contributes to folate deficiency and is particularly problematic when associated with alcohol abuse and smoking. In the UK alone, it is estimated that over 13 million people currently consume too little folates in their diet (Food Standards Agency, 2007).

Malabsorption, intestinal disease, liver or renal failure and use of certain medications (valproic acid, carbamazepine, trimethoprim, phenobarbitol, primidone, diphenylhydantoin, oxcarbamazepine, sulfonamides and methotrexate) are other conditions where folate requirements are increased. Presence of any of these factors in addition to inadequate absorption of food folate and further losses through cooking practices, means that considerable number of women of reproductive age fail to get enough daily folates from diet alone.

5. Effects of folate deficiency in mother

Folate deficiency is known to lead to megaloblastic anemia, which is characterized by large immature RBCs in the peripheral smear. Early symptoms of such deficiency are non-specific and may include tiredness, irritability and loss of appetite. As anaemia worsens symptoms of fatigue, shortness of breath, headache, gastrointestinal upsets and palpitations may appear. If left untreated, megaloblastic anaemia may be fatal. Diagnosis of a subclinical deficiency relies on demonstrating reduced red cell folate concentration or on other biochemical evidence such as increased homocysteine concentration, since haematological manifestations are usually absent. In advanced cases, the presence of oval macrocytes and hyper segmented neutrophils in the peripheral smear is highly sensitive and specific for the diagnosis of megaloblastic anaemia. Folate deficiency during pregnancy has also been linked to preterm birth, low infant birth weight, intrauterine foetal growth restriction and placental abruption, although various studies have reported conflicting results regarding these complications (Martí-Carvajal *et al*., 2004; Nilsen *et al*., 2008; Scholl & Johnson, 2000; Ananth *et al.,* 2008; Ananth *et al.,* 2007).

6. Folate deficiency and birth defects

Maternal folate deficiency is associated with fetal congenital malformations. The best documented are neural tube defects (NTDs), including anencephaly and spina bifida. Each year these two most common forms of neural-tube defects occur in 1 in 1000 pregnancies in

the United States (Cragan *et al*., 1995) and in an estimated 300,000 or more new-borns worldwide (Shibuya & Murray, 1998). In the UK there are estimates of between 700 and 900 pregnancies with neural tube defects each year.

There are a few studies suggesting that folate deficiency may also lead to other birth defects including cleft lip, cleft palate, certain heart defects and limb malformations (Maldonado *et al*., 2011; Salerno *et al*., 2008). However, a recent Cochrane review concluded that folic acid, alone or in combination with vitamins and minerals decreases NTD risk but did not have a clear effect on the occurrence of other birth defects (De-Regil *et al*., 2010).

7. Other effects of folate deficiency

Studies have associated low folate levels with an increased risk of Alzheimer´s disease and depression, although robust evidence for such an association is still lacking and results from further studies are eagerly awaited (Malouf & Grimley, 2008; Kronenberg *et al*., 2009; Stanger *et al*., 2009). Numerous studies have also indicated that elevated levels of homocysteine in the blood directly or indirectly increase the risk of atherosclerosis, coronary heart disease and stroke. By virtue of its involvement in the metabolic pathways, deficiency of folate can increase homocysteine levels. Low folate levels therefore have been described as a risk for cardiovascular disease by some authors (Voutilainen *et al*., 2001) and folate supplementation has been shown to decrease homocysteine levels and to improve endothelial function (Doshi *et al*., 2002; Doshi *et al*., 2001;Wald *et al*., 2001a). However due to uncertain nature of the existing evidence and conflicting reports in the literature, several randomised controlled trials are being conducted to establish whether folic acid supplementation to reduce homocysteine levels can lower the actual rates of cardiovascular diseases. A meta-analysis of 8 large, randomized, placebo-controlled trials of folic acid supplementation involving 37,485 individuals at increased risk of cardiovascular disease demonstrated that although the folic acid allocation yielded an average 25% reduction in homocysteine levels, the supplementation had no significant effects within 5 years on cardiovascular events or on overall cancer or mortality in the populations studied (Clarke *et al*., 2010).

8. Geographic variations in folate deficiency

Folate deficiency is very common in many parts of the developing world and is a part of the wider problem of poverty and malnutrition. Although much less prevalent, in developed countries, folate deficiency may be encountered mostly in socially or economically underprivileged groups or those with specific medical conditions.

9. Folic acid toxicity

The risk of toxicity from folic acid is low, because folate is a water-soluble vitamin and easily excreted by the kidneys. It has been suggested that high levels of folic acid can provoke seizures in some patients taking anti-convulsant medications (Herbert, 1999). There have also been concerns about the masking of vitamin B_{12} deficiency by folic acid supplementation. Vitamin B_{12} deficiency has been estimated to affect up to 10–15% of the population over 60 years of age. Folic acid supplements can correct the anemia associated with vitamin B_{12} deficiency, but they do not correct the changes in the nervous system that result from vitamin B_{12} deficiency. Permanent nerve damage could theoretically occur due to such masking. The elderly population and those with folate intake above 1000 micrograms per day are at particular risk. To date the evidence that such masking actually occurs is scarce, and studies indicate that food fortification has not caused a major increase in masking of vitamin B_{12} deficiency (Mills *et al*., 2003). It is important, however, for both

elderly individuals and health professionals to be aware of this relationship between folic acid and vitamin B_{12}. In most countries including the US, health authorities have set the tolerable upper intake level of folic acid from fortified foods or supplements at 1,000 μg/day for adults.

10. Folic acid and risk of colorectal cancer and adenoma

Various experimental and observational studies have raised concerns about a possible role of folic acid in the development of colorectal carcinomas and adenomas. It has been suggested that folic acid may have dual modulatory effects on carcinogenesis depending on the timing and dose of its administration (Kim, 2003).

Animal studies have pointed to a possibility that increased folic acid intake may decrease the risk of early neoplastic lesions, but that supplementation in higher amounts once a neoplasm has developed may actually increase the risk of colorectal cancer (Song *et al*., 2009a; Song *et al*., 2009b). However, extrapolation of observations from animal models to humans should be undertaken with caution. The folic acid fortification began in the US and Canada in 1996 and 1997 respectively and concurrently both countries experienced an increase in rates of colorectal cancers which lead to the hypothesis that the institution of folic acid fortification may be wholly or partly responsible for the increase in colorectal cancer rates (Mason *et al*., 2007).

Although the bulk of the evidence to date favours a good safety profile for folic acid supplements, the present scientific consensus on the association of folic acid with colon cancer appears to be divided. Many early epidemiological studies demonstrated a protective effect of folic acid supplements against colorectal and other organ malignancies, while some more recent studies have had conflicting results (Kim, 2004a; McCullough & Giovannucci, 2004; Crott & Mason, 2005; Choi & Mason, 2002; Farber, 1949; Heinle & Welch, 1948; Cole *et al*., 2007; Mason *et al*., 2008) with some investigators arguing for and others against folic acid fortification.

A meta-analysis challenged the results from epidemiological studies which indicated that folic acid status is inversely related to the risk of developing colorectal tumours. The results suggested an increased long-term risk of colorectal adenoma, especially advanced adenoma, among those participants randomised to the folic acid supplementation group (Fife *et al*., 2011). Another systematic review suggested that higher folic acid intake levels reduced the risk of developing colorectal cancer (Kennedy *et al*., 2011). A pooled analysis of 13 prospective cohort studies found that a higher folic acid intake is modestly associated with a reduced risk of colon cancer. A 2% risk reduction (95% CI 0-3%) was estimated for every 100 μg/day increase in total folic acid intake (Kim *et al*., 2010). Another prospective study involving use of validated food-frequency questionnaires examined the latency between folic acid intake and the incidence of colorectal cancer. The authors did not find any clear evidence that the intake of folic acid supplements increased risk of colorectal cancer. They in fact supported the hypothesis that folic acid intake decreased the risk of initiation or early development of colorectal cancer (Lee *et al*., 2011).

As is evident from the conclusions derived from various studies, the question as to whether folic acid plays a role in colon cancer reveals mixed results at the present time. A number of factors need to be considered when evaluating and comparing studies, including limitations in the current literature. There is a great deal of heterogeneity in the populations studied, including geographical and ethnic differences, presence of polymorphisms, use of data from both pre- and post-fortification time periods,

and inclusion of populations with and without fortification. Different methods and cut-offs are used to categorize highest compared to lowest intake and there is variability in serum levels, with a single serum level used to evaluate risk. The length of time for follow up varies between studies and may not be adequate to evaluate impact on risk. This makes it difficult to form a single conclusion regarding the impact of folate on cancer risk. However, there are populations who appear to be at risk due to either inadequate or excess intake.

As more evidence regarding folic acid and colon cancer emerges in the future, if it is established that folic acid supplements do play a part in promoting tumorigenesis, the implications of such findings would be very significant from public health point of view. It would necessitate a reconsideration of the mandatory food fortification programs already in place in various countries. Long-term follow-up studies in countries that have adopted mandatory folic acid fortification are therefore the need of the hour and their results will likely determine the fate of folic acid fortification programs in the future. In the meantime, the effects of folic acid fortification in the population will need careful monitoring (Kim, 2004b).

11. Recommendations for folic acid supplementation

11.1 Prevention of neural tube defects

The suggestion that folate deficiency might play a part in the aetiology of NTDs was made originally by Brian Hibbard in 1964 (Hibbard, 1964). It is vital for women to have adequate folate stores before conception as neural tube development occurs in the first few weeks after conception, typically before the woman is aware of her pregnancy. Peri-conceptional use of folic acid supplements reduces the risk of the first occurrence, as well as the recurrence of NTDs (relative risk (RR) 0.28, 95% confidence interval (CI) 0.13–0.58) (Lumley *et al*., 2001). The exact mechanism how folic acid prevents birth defects is unknown however differential methylation of the insulin-like growth factor 2 (IGF-2) gene or abnormal homocysteine metabolism are some of the possibilities suggested (Steegers-Theunissen *et al*., 2009). Supplementation with folic acid has also been shown to reduce the risk of structural cardiac and craniofacial abnormalities (Finnell *et al*., 2004). The level of protection offered by folic acid against NTDs increases with the dosage.

Serum folate concentrations increase by 0.94 ng/mL for every 0.1 mg/day increase in folic acid intake in women aged 20-35 years, and about double that in people aged 40-65. Every doubling of serum folate concentration roughly halves the risk of a NTD. An increase of 0.4 mg/day would reduce risk by about 36%, of 1 mg/day by 57%, and taking a 5 mg tablet daily would reduce risk by about 85% (Wald *et al*., 2001b). In women at high risk of foetal NTDs (due to previous pregnancy with NTD), the MRC randomised double-blind prevention trial has shown that a higher dose of folic acid supplementation (4 mg/day) reduces the risk of a subsequent NTD-affected pregnancy by 72% (RR 0.28, 95% CI 0.12–0.71) (MRC Vitamin Study Research Group, 1991).

Based on the current evidence it is therefore recommended that all women of child bearing age should take 0·4 mg (400 micrograms) of folic acid daily when planning a pregnancy. Those women who have had a previous pregnancy affected by a neural-tube defect should take 5 mg folic acid daily periconceptionally, starting at least one month before conception and continuing throughout the first trimester of pregnancy.

11.2 Obesity

Over the years data have emerged that obese women are at an increased risk of neural tube defects in their offspring and require a higher dose of peri-conceptional folic acid supplementation. A meta-analysis of 12

observational cohort studies reported an odds ratio (OR) of 1.22 (95% CI 0.99–1.49), 1.70 (95% CI 1.34–2.15) and 3.11 (95% CI 1.75–5.46) for women defined as overweight, obese and severely obese, respectively, compared with healthy-weight women (Rasmussen *et al*., 2008). There is also evidence from cross-sectional data that, compared to women with a BMI <27, women with a BMI more than 27 are less likely to use nutritional supplements and less likely to receive folate through their diet. Even after controlling for folate intake, women with a BMI >27 have lower serum folate levels compared to women with a BMI <27 (Mojtabai, 2004). There is a higher risk of NTDs associated with increased maternal weight, even after universal folic acid flour fortification (Ray *et al*., 2005). Based on these data, the 2010 CMACE/RCOG Joint Guidelines for Management of Women with Obesity in Pregnancy recommend that women with a BMI more than 30 wishing to become pregnant should be advised to take 5 mg folic acid supplementation daily starting at least one month before conception and continuing during the first trimester of pregnancy (Modder & Fitzsimons, 2010).

11.3 Diabetes and Epilepsy

Diabetes and epilepsy are other conditions that confer an increased risk of NTDs in pregnancy. It is recommended that women with these conditions consume a larger dose (5mg) of folic acid. **Table 1** summarizes the conditions where higher dose of folic acid supplementation is recommended in pregnancy. After the first trimester, most women receiving a higher dose of folic acid can be advised to transition to a daily multivitamin supplement containing 0.4 mg of folic acid for the duration of pregnancy and lactation.

Some authors have suggested a role for measurement of red blood cell (RBC) folate concentration in order to determine the most appropriate dose of folic acid supplements, and this may be useful in patients with medical conditions such as diabetes, obesity or epilepsy (Tam et al., 2009). It has been shown that RBC folate concentrations greater than 906 nmol/L are maximally protective against folate-dependent NTDs (Daly et al., 1995).

Table 1: : *Conditions in which a higher dose (5mg) of peri-conceptional folic acid is recommended*

A)	Clinical conditions associated with decreased folate levels • Alcohol abuse • Malabsorption/ intestinal disease or gastric bypass • Haemolytic anaemias • Liver disease • Renal failure • Medications (alone or in combination) - valproic acid, carbamazepine, trimethoprim, phenobarbitol, primidone, diphenylhydantoin, oxcarbamazepine, sulfonamides and methotrexate.
B)	Clinical conditions associated with an increased risk of fetal NTDs • Personal or family history of NTDs • Obesity • Pre-gestational Diabetes • Epilepsy

11.4 Megaloblastic anaemia

In folate deficiency megaloblastic anaemia, daily folic acid supplementation (5 – 15 mg in mal absorption states) for 4 months brings about haematological remission and replenishes the body stores. However, in managing acute cases of megaloblastic anaemia, the treatment often has to be started before a diagnosis of the specific cause (vitamin B_{12} or folate deficiency) has been made. To avoid complications that may arise due to masking of B_{12} deficiency with folic acid, supplementation with both folic acid and vitamin B_{12} is recommended until specific diagnosis is available.

12. Summary

Folic acid has been a long companion not only of the obstetricians but also the primary care physicians and general practitioners in the peri-conceptional care of women in reproductive age group. With the emergence of new evidence of varying requirements of folic acid in different population subgroups it has become important for healthcare professionals to be aware of varying needs and recommend the correct dose of supplements to their patients.

In many populations across the world, awareness of the benefits of folic acid before conception and during pregnancy is low. Health care professionals, magazines, newspapers, radio and television are the common sources of information. Effective knowledge translation is needed to ensure that women are informed about the benefits of folic acid during the reproductive years. This knowledge will allow them to make informed decisions about folic acid consumption. Health care professionals play an influential role in promoting folic acid knowledge among women of childbearing age. Lower levels of knowledge among women with lower levels of education and/or household income must be addressed.

Although recent years have seen concerns raised over the safety of folic acid supplementation, the majority of the evidence existing to date appears reassuring. Long term population follow up studies as well as rigorous laboratory research are likely to better define the role of folic acid in the development of colorectal tumors in the future.

Funding

No funding or support of any kind was received for the preparation of this manuscript.

References

MRC Vitamin Study Research Group. (1991). Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. *Lancet, 338*(8760), 131-137.

Ananth, C.V., Peltier, M.R., De Marco, C., Elsasser, D. Getahun, D., Rozen, R. & Smulian, J.C. (2007). Associations between 2 polymorphisms in the methylenetetrahydrofolate reductase gene and placental abruption. *American Journal of Obstetrics and Gynecology, 197*(4), 385.e1-7.

Ananth, C.V., Peltier, M.R., Moore, D.F., Kinzler, W.L., Leclerc, D. & Rozen, R.R. (2008). Reduced folate carrier 80A-->G polymorphism, plasma folate, and risk of placental abruption. *Human Genetics, 124*(2), 137-145.

Choi S.W. & Mason J.B. (2002). Folate status: effects on pathways to colorectal carcinogenesis. *132*(8 supp.), 2413S – 2418S.

Clarke R, Halsey J, Lewington S., Lonn, E., Armitage, J., Manson, J.E. Bonaa, K.H., Spence, J. D., Nygart, O., Jamison, R., Gaziano, J.M., Guarino, P. Bennett, D., Mir, F., Peto, R. & Collins, R. (2010). B-Vitamin Treatment Trialists' Collaboration. Effects of lowering homocysteine levels with B vitamins on cardiovascular disease, cancer, and cause-specific mortality: Meta-analysis of 8 randomized trials involving 37,485 individuals. *Archives of internal Medicine, 170*(18), 1622-1631.

Cole, B.F., Baron, J.A., Sandler, R.S., Haile, R.W., Ahnen, D.J., Bresalier, R.S., McKeown-Eyssen, G., Summers, R.W., Rothstein, R.I., Vurke, C.A., Snover, D.C., Church, T.R., Allen, J.I., Robertson, D.J., Beck, G.J., Bond, J.H., Byers, T. [et al]. (2007). Folic acid for the prevention of colorectal adenomas: a

randomized clinical trial. *JAMA, 297*(21), 2351 – 2359.

Cragan, J.D., Roberts, H.E., Edmonds, L.D.,Khoury, M.J., Kirvy, R.S., Shaw, G.M., Velie, E.M., Merz, R.D., Forrester, M.B., Williamson, R.A., Krishnamurthi, D.S., Stevenson, R.E. & Dean, J.H. (1995). Surveillance for anencephaly and spina bifida and the impact of prenatal diagnosis -- United States, 1985-1994. *Morbidity and Mortality Weekly Report. CDC Surveill Summaries, 44*(4), 1-13.

Crott, J. & Mason J.B. (2005). MTHFR polymorphisms and colorectal neoplasia. In P. Ueland, & R. Rozen (eds.), *MTHFR polymorphisms and disease* (pp. 179 – 184.). Georgetown (TX): Landes Bioscience.

Daly, L.E., Kirke, P.N., Molloy, A., Weir, D.G. & Scott, J.M. (1995). Folate levels and neural tube defects. Implications for prevention. *JAMA, 274*(21), 1698-1702.

De-Regil, L.M., Fernández-Gaxiola, A.C., Dowswell, T. & Pena-Rosas, J.P. (2010). Effects and safety of periconceptional folate supplementation for preventing birth defects. *The Cochrane Database Systermatic Reviews, 10: CD007950*. doi:10.1002/14651858.CD007950.pub2

Doshi, S.N, McDowell, I.F., Moat S.J., Lang, D., Mewcombe, R.G., Kredan, M.B., Lewis, M.J. & Goodfellow, J. (2001). Folate improves endothelial function in coronary artery disease: an effect mediated by reduction of intracellular supperoxide? *Arteriosclerosis, Thrombosis & Vascular Biology, 21*(7), 1196-1202.

Doshi, S.N, McDowell, I.F., Moat, S.J., Payne, M., Durrant, H.J., Lewis, M.J & GoodFellow, J. (2002). Folic acid improves endothelial function in coronary artery disease via mechanisms largely independent of homocysteine Lowering. *Circulation, 105*(1), 22-26.

Farber, S. (1949). Some observations on the effect of folic acid antagonists on acute leukemia and other forms of incurable cancer. *Blood, 4*(2), 160 – 167.

Fife, J., Raniga, S., Hider, P.N. & Frizelle, F.A. (2011). Folic acid supplementation and colorectal cancer risk: a meta-analysis. *Colorectal Disease, 13*(2), 132-137. doi:10.1111/j.1463-1318.2009.02089.x.

Finnell, R.H., Shaw, G.M., Lammer, E.J., Brandl, K.L., Carmichael, S.L. & Rosenquist, T.H. (2004). Gene-nutrient interactions: importance of folates and retinoids during early embryogenesis. *Toxicology and applied Pharmacology, 198*(2), 75-85.

Food and Drug Administration. (1996). Food standards: amendment of standards of identity for enriched grain products to require addition of folic acid; correction. 21 CFR Parts 136, 137, and 139. *Federal Register, 61*(151), 40513.

Food Standards Agency. (2007). *Board recommends mandatory fortification.* Food Standards Agency. Retrieved from http://www.food.gov.uk/news/newsarchive/2007/may/folatefort

Heinle, R.W., & Welch, A.D. (1948). Experiments with pteroylglutamic acid and pteroylglutamic acid deficiency in human leukemia [abstract]. *Journal of Clinical Investication, 27*(4), 539.

Herbert, V. (1999). Folic Acid. In M. Shils, J.Olson, M. Shike, & A.C. ROss (eds.), *Nutrition in Health and Disease.* Baltimore: Williams & Wilkins.

Hibbard, B.M. (1964). The role of folic acid in pregnancy with particular reference to anaemia, abruption and abortion. *Journal of Obstetrics and Gynaecology British Communication, 71*, 529-542.

Kennedy, D.A., Stern, S.J., Matok, I., Sarkar, M., Nickle, C. & Koren, G. (2011). Folate intake and the risk of colorectal cancer: a systematic review and meta-analysis. *Cancer Epidemiology, 35*(1), 2-10. doi:10.1016/j.canep.2010.11.004.

Kim, D.H., Smith-Warner, S.A., Spiegelman, D.,, Yaun, S.S., Colditz, G.A., Freudenheim, J.L.,

Giovannucci, E., Goldbohm, R.A., Graham, S., Harnack, L., Jacobs, E.J., Leitzmann, M., Mannisto, S., Miller, A.B., Potter, J.D., Rohan, T.E., e.t al. (2010). Pooled analyses of 13 prospective cohort studies on folate intake and colon cancer. *Cancer Causes Control, 21*(11), 1919-1930. doi:10.1007/s10552-010-9620-8.

Kim, Y.I. (2003). Role of folate in colon cancer development and progression. *The Journal of Nutrition, 133*(11 supp.1), 3731S–3739S.

Kim, Y.I. (2004a). Folate and DNA methylation: a mechanistic link between folate deficiency and colorectal cancer? *Cancer Epidemiology Biomarkers and Prevention, 13*(4), 511 – 519.

Kim, Y.I. (2004b). Will mandatory folic acid fortification prevent or promote cancer? *American Journal of Clinical Nutrition, 80*(5), 1123–1128.

Kronenberg, G., Colla, M. & Endres, M. (2009). Folic acid, neurodegenerative and neuropsychiatric disease. *Current Moluculor Medicine, 9*(3), 315-323.

Lee, J.E., Willett, W.C., Fuchs, C.S., Smith-Warner, S.A., Wu, K., Ma, J. & Giovannucci, E. (2011). Folate intake and risk of colorectal cancer and adenoma: modification by time. *The American Journal of Clinical Nutrition, 93*(4), 817-825. doi:10.3945/ajcn.110.007781

Lumley, J., Watson, L., Watson, M. & Bower, C. (2001). Periconceptional supplementation with folate and/or multivitamins for preventing neural tube defects. *Cochrane Database of Systematic Reviews, 3*.

Maldonado, E., Murillo, J., Barrio, C.,del Rio, A., Perez-miguelsanz,J.,Lopez- Gordillo, Y.,Partearroyo, T., Paradas, I., Maestro, C., Martinez- Samz, E., Garela- Moreiras, G. & Martinez- alvarez, C. (2011). Occurrence of Cleft-Palate and Alteration of Tgf-β(3) Expression and the Mechanisms Leading to Palatal Fusion in Mice following Dietary Folic-Acid Deficiency. *Cells, Tissues, Organs, 194*(5), 406-420.

Malouf, R. & Grimley Evans, J. (2008). Folic acid with or without vitamin B12 for the prevention and treatment of healthy elderly and demented people. *Cochrane Database System Review, 4: CD004514*. doi:10.1002/14651858.CD004514

Martí-Carvajal, A., Peña-Martí, G., Comunián-Carrasco, G., munoz-navarro, S., luco, M., Marti- pena, A. & Medina- Laurentin, C. (2004). Prematurity and maternal folate deficiency: anemia during pregnancy study group results in Valencia, Venezuela. *Archivos Latinoamericanos de Nutricion, 54*(1), 45-49.

Mason, J.B., Cole, B.F., Baron, J.A., Kim, Y.I. & Smith, A.D. (2008). Folic acid fortification and cancer risk. *Lancet, 371*(9621), 1335-1336. doi:10.1016/S0140-6736(08)60590-5

Mason, J.B., Dickstein, A., Jacques,P.F., Haggarty, P., Sclhuv, J. Dallal, G. & Rosenberg,I.H.et al. (2007). A Temporal Association between Folic Acid Fortification and an Increase in Colorectal Cancer Rates May Be Illuminating Important Biological Principles: A Hypothesis. *Cancer Epidemiology Biomarkers and Prevention, 16*(7), 1325-1329.

McCullough, M.L. & Giovannucci, E.L. (2004). Diet and cancer prevention. *Oncogene, 23*(38), 6349 – 6364.

Mills, J.L., Von Kohorn, I., Conley, M.R., Zeller, J.A., Cox, C., Williamson, R.E. & Dufour, D. R. (2003). Low vitamin B-12 concentrations in patients without anemia: the effect of folic acid fortification of grain. *The American Journal of Clinical Nutrition., 77*(6), 1474-1477.

Mitchell, H.K., Snell, E.E. & Williams, R.J. (1941). The concentration of "Folic acid". *Journal of American Chemical Society, 63*(8), 2284.

Modder, J. & Fitzsimons, K.J. (2010). *CMACE/RCOG Joint Guideline: Management of Women with Obesity in Pregnancy.* England: Centre for Maternal and Child Enquiries & Royal College of Obstetricians and Gynaecologists.

Mojtabai, R. (2004). Body mass index and serum folate in childbearing age women. *European Journal of Epidemiology, 19*(11), 1029-1036.

Molloy, A.M., Daly, S., Mills, J.L., Kirke, P.N., White head, A.S., Rams bottom, D., Conley, M.R., Weir, D.G. & Seott, J.M. (1997). Thermolabile variant of 5,10-methylenetetrahydrofolate reductase associated with low red-cell folates: implications for folate intake recommendations. *Lancet, 349*(9065), 1591-1593.

Nilsen, R.M., Vollset, S.E., Rasmussen, S.A., Ueland, P.M. & Daltveit, A.K. (2008). Folic acid and multivitamin supplement use and risk of placental abruption: a population-based registry study. *American Journal of Epidemiology, 167*(7), 867-874.

Obican, S.G., Finnell, R.H., Mills, J.L., Shaw, G.M. & Scialli, A.R. (2010). Folic acid in early pregnancy: a public health success story. *FASEB Journal, 24*(11), 4167-4174.

Rasmussen, S.A., Chu S.Y., Kim, S.Y., Schmid,C.H. & lau, J. (2008). Maternal obesity and risk of neural tube defects: a metaanalysis. *American Journal of Obstetrics and Gynecology, 198*(6), 611-619.

Ray, J.G., Wyatt, P.R., Vermeulen, M.J., Meier, C. & Cole, D.E. (2005). Greater Maternal Weight and the Ongoing Risk of Neural Tube Defects After Folic Acid Flour Fortification. *Obstetrics & Gynecology, 105*(2), 261-265.

Salerno, P., Bianchi, F., Pierini, A., Baldi, F., Carbone, P., Mantovani, A. & Taruscio, D. (2008). Folic acid and congenital malformation: scientific evidence and public health strategies. *Annali di Igiene, 20*(6), 519-530.

Scholl, T.O., Johnson, W.G. (2000). Folic acid: influence on the outcome of pregnancy. *The American Journal of Clinical Nutrition, 71*(5-Supp.), 1295S-1303S.

Shibuya, K. & Murray, C.J.L. (1998). Congenital anomalies. In C.J.L.Murray, & A.D.Lopez (eds.), *Health dimensions of sex and reproduction: the global burden of sexually transmitted diseases, HIV, maternal conditions, perinatal disorders, and congenital anomalies* (pp. 455-512). Boston: Harvard School of Public Health.

Song, J., Sohn, K.J., Medline, A., Ash, C., Gallinger, S. & Kim, Y.I. (2000b). Chemopreventive effects of dietary folate on intestinal polyps in Apc +/2Msh22/2 mice. *Cancer Research, 60*(12), 3191–3199.

Song, J., Medline, A., Mason, J.B., Gallinger, S. & Kim, Y.I. (2000a). Effects of dietary folate on intestinal tumorigenesis in the apcMin mouse. *Cancer Research, 60*(19), 5434–5440.

Stanger, O., Fowler, B., Piertzik, K., Huemer, M., Hasehke- vecher, E., Semmler, A., Lorenzl, S. & Linnevanka, M. (2009). Homocysteine, folate and vitamin B_{12} in neuropsychiatric diseases: review and treatment recommendations. *Expert Review of Neurotherapeutics, 9*(9), 1393-1412.

Steegers-Theunissen, R.P., Obermann-Borst, S.A., Kremer, D., Lindemans, J., Siebel, C., Steegers, E.A., Slagboom, P.E. & Heijmans, B.T. (2009). Periconceptional maternal folic acid use of 400 microg per day is related to increased methylation of the IGF2 gene in the very young child. *PLoS One, 4*(11), e7845. doi:10.1371/journal.pone.0007845.

Tam, C., McKenna, K., Goh, Y.I, Klieger-Grossman, C., O'Connor, D.L., Einarson, A. & Koren, G. (2009). Periconceptional folic acid supplementation: A new indication for therapeutic drug monitoring. *Ther Drug Monit, 31*(3), 319-326. doi: 10.1097/FTD.0b013e31819f3340.

Voutilainen, S., Rissanen, T.H., Virtanen, J., Lakka, T.A. & Salonen, J.T. (2001a). Low dietary folate intake is associated with an excess incidence of acute coronary events: The kuopio ischemic heart disease risk factor study. *Circulation, 103*(22), 2674-2680.

Wald, D.S., Bishop, L., Wald, N.J., Law, M., Hennessy, E., Weir, D., McPartlin, J. & Scott, J. (2001a). Randomized trial of folic acid

supplementation and serum homocysteine levels. *Archives of Internal Medicine, 161*(5), 695-700.

Wald, N.J., Law, M.R., Morris, J.K. & Wald, D.S. (2001b). Quantifying the effect of folic acid. *Lancet, 358*(9298), 2069-2073.

Wills, L. (1931). Treatment of "pernicious anaemia of pregnancy" and "tropical anaemia" . *British Medical Journal, 1*(3676), 1059-1064.

Review Article

Pesticide and Microbial Contaminants of Groundwater and their Removal Methods: A Mini Review

[1,2]Julia Plattner, [1,2]Christian Kazner, [1]Gayathri Naidu, [2]Thomas Wintgens, [1]Saravanamuthu Vigneswaran*

[1]*Faculty of Engineering, University of Technology Sydney (UTS), Australia*
[2]*University of Applied Sciences and Arts North-western Switzerland School of Life Sciences, Institute for Ecopreneurship, Switzerland*

Received: 5 December 2017 , Accepted: 1 May 2018

Abstract:

The world's increasing population, economic development and climate change are driving the demand for more drinking water. In India more than 100 million people live in areas of poor water quality. It has been reported that more than 33% of India's groundwater resources are unsuitable for consumption due to pesticide and microbe contamination. This paper discusses the types of pesticides and microbes predominantly present in groundwater in India. Details on the environmental and human health implication of consuming groundwater containing these contaminants are highlighted. Physiochemical properties of pesticides and microbes such as the molecular weight, hydrophobicity and polarity are evaluated in detail. These information provides useful understanding on the fundamental causes of widespread pesticides and microbes contamination in groundwater. Technologies for removal of pesticides and microbes are discussed. The advantages and disadvantages of each technology for effectively removing pesticides and microbes to attain portable water standard is highlighted. The effectiveness of a technology for removing pesticides and microbes is dependent on the physiochemical properties of the contaminants.

Keywords: Contamination, Drinking water, Ground water, Microbes, Pesticides

1. Introduction

The intensive use of pesticides has led to widespread contamination of the biotic as well as the abiotic environment. Pesticides are found in surface water and also in groundwater sources all over the world.

This paper discusses the pesticides and microbes present in groundwater in India. Groundwater in India is largely used for domestic purposes by 80% of the rural population and 50% of the urban population. In many areas, groundwater serves as the only source of drinking water (Chakraborti *et al.*, 2011). Over abstraction of groundwater and river water for agricultural purposes leads to rapidly dropping water tables. Untreated sewage flowing in open drains and open landfills with no protection from leaching into the groundwater are the main man-made sources for water contamination (Central Pollution Control Board, 2008). It has been reported that more than 33% of India's groundwater resources are unsuitable for consumption (Chakraborti *et al.*, 2011).

Corresponding author: Dr. SaravanamuthuVigneswaran , , Faculty of Engineering, University of Technology Sydney (UTS), P.O. Box 123, Broadway, NSW 2007 Australia. email: Saravanamuth.Vigneswaran@uts.edu.au

Anthropogenic pollution such as microbial contaminants, nitrate, pesticides and industrial discharge, together with geogenic contaminants such as fluoride, arsenic, iron and saline water, pose a threat to human health (McFarlane & Williamson, 2002).

2. Microbial and chemical contaminants in groundwater in India

2.1. Microbial contamination

It is estimated that currently only 10% of sewage generated in India's cities is treated prior to reaching groundwater or surface water resources (Chaudhary *et al*., 2002). The main problem resulting from this is the microbial contamination of drinking water. Microbes such as bacteria, protozoa and viruses are pathogens, which can cause diseases.

Bacteria: Bacterial contamination can be removed largely through soil passage (Ayuso-Gabella *et al*., 2011; Tielemans, 2007). Therefore bacterial contamination applies mainly to surface water rather than groundwater. However, due to poor sanitation practice and infrastructure, bacterial contamination can also be found in shallow tube wells with a depth of around 10 m. With increasing depth, the contamination of bacteria decreases (Chakraborti *et al*., 2011). Another source of bacterial contamination is the handling of the collected water in rural areas. Since many households have no direct fresh water access, water has to be collected daily and is stored in the house for use. This practice, in addition to the already contaminated collected water, increases the risk of microbial growth, especially in hot climate. If water contaminated with bacteria is consumed, the most common consequence is diarrhoea caused by faecal coliform bacteria, such as *Escheria Coli 0157:h7* or *Enterococci*. *Enterococci* are specifically a risk for people with a weak immune system such as elderly, young children and infants. *Legionella* is a bacterium found naturally in the environment. If it is aerosolized for example in a shower or during air conditioning and inhaled, legionnaires disease, a type of pneumonia, can be caused (Woodie, 2014). *Salmonella* can cause intestinal illness as well as typhoid and paratyphoid fever (Levantesi *et al*., 2012).

Viruses: Viruses present another major group of microbial contamination causing waterborne diseases such as hepatitis A, polio, meningitis, fever and gastroenteritis. Viruses are very small (10-100 nm) and often resistant for disinfection. Filtration followed by disinfection has the best removal efficiency as viruses are often occurring in colloids with organic matter. In India viruses pose a high threat to the population; therefore the Indian Drinking Water Specification (BIS 2012 IS 10500:2012) states that the water should be free from viruses.

Protozoa: Contamination with *Gardialamblia* and *Cryptosporium* are one of the main microbial hazards in drinking water. Symptoms of infection are diarrhoea, abdominal cramps, headache, and blood and weight loss. The symptoms usually start from two to twenty five days after swallowing the cyst (Vermont Department of Health, 2015).

Pathogens can be largely inactivated through disinfection with chlorination tablets, ozone, heat or UV-light, or removed with membrane filtration technologies (Ultrafiltration/ Nanofiltration/ Reverse Osmosis). For drinking water application, it is advised to apply a multi barrier approach to establish and maintain disinfection (Bennett, 2008).

2.2. Pesticide contamination

The intensive use of pesticides in India has led to widespread contamination of the biotic as well as the abiotic environment in India (Yadav *et al*., 2015). Pesticides were found in surface water and groundwater (Lari *et al*., 2014; Mutiyar *et al*., 2011; Sankararamakrishnan *et*

al., 2005; Sikarwar *et al*., 2014). In 2005, Singh and his co-workers estimated that 9,000 tons of pesticides may enter into the river system each year. Globally up to 3 million cases of acute, severe poisoning with pesticides have been recorded. Pesticide poisoning is more frequent in developing countries than in industrialized countries, even though only 25% of the pesticides are consumed in developing countries. The reasons for increased risk of pesticide poisoning in developing and transitional countries are low level of protection of workers due to inadequate availability of protective equipment, improper application, condition and practices, unsafe storage and disposal facilities, poor management, untrained pesticide dealers and inadequate health centres (Yadav *et al*., 2015). A multitude of pesticides have been identified as endocrine disrupting chemicals (EDC). In 2011, Mnif and his co-workers have compiled an overview of 105 substances and their effects on the hormone system of humans. The combined effects of pesticides are a major health risk to humans. But, wildlife is also particularly vulnerable to the toxic and endocrine effects of pesticides.

In India, 76% of the used pesticides are accounted for as insecticides, 13% as fungicides and 10% as herbicides. In cotton and paddy cultivation, more than 50% of the pesticides are used. Depending on the climatic conditions and the agriculture in different states in India, the pesticide consumption can vary strongly from one state to another. Uttar Pradesh, Maharashtra and Andhra Pradesh are the states with the highest pesticide consumption (Yadav *et al.*, 2015). Based on the precautionary principle, the European Drinking Water Directive (98/83/EC) has set a limit of 0.1 µg/L for single pesticides with an exception of aldrin, dieldrin, heptachlor and heptachlorepoxide (where the limit for the single compound is 0.03 µg/L) and 0.5 µg/L for the sum of all active pesticides detected. India has set individual values for 18 substances in the Indian Standard Drinking Water Specification as shown in the table below (**Table** 1). It can be seen that the European standards are much stricter and less compound specific than the Indian threshold values.

In general, the mobility of pesticides and thus their risk of leachability into the groundwater have been correlated with a weak adsorption of the soil matrix quantified in terms of a small soil organic carbon-water partitioning coefficient (K_{OC}) (Arias-Estévez *et al*., 2008). Generally, pesticides with $K_{OC} \leqslant 1{,}000$ are potentially leaching compounds whereas pesticides with $K_{OC} \geqslant 1{,}000$ have also been found in the groundwater. This is mainly due to the fact that when the pesticides are polar, their behaviour does not follow the rule. Also, the specific site and application along with the soil type and the climatic conditions play an important role on the fate of the pesticide in the environment (Arias-Estévez *et al*., 2008).

Table 1: *Pesticide residue limits according to the Indian drinking water guidelines [Indian Standard Drinking Water Specification (second revision) 2012]*

Pesticide	Limit [µg/L]
Alachlor	20
Atrazine	2
Aldrin/Dieldrin	0.03
Alpha HCH	0.01
Beta HCH	0.04
Butachlor	125
Chlorpyriphos	30
Delta HCH	0.04
2,4- Dichlorophenoxyacetic acid	30
DDT (*o,p*- and *p,p*- isomers of DDT, DDE and DDD)	1
Endosulfan (*alpha* and *beta* sulphates)	0.4
Ethion	3
Gamma HCH (Lindane)	2
Isoproturon	9
Malathion	190
Methyl parathion	0.3
Monocrotophos	1
Phorate	2

Another important chemical characteristic of pesticides that plays a role in their behaviour in the environment or in a treatment system is how hydrophilic or hydrophobic they are and hence, how prone they are to stay in the water phase or to get adsorbed in a different material or fluid. The partition coefficient (P) describes this behaviour using the logarithm of the ratio (Log P) of a compound in a polar (e.g. water) and non-polar solvents (e.g. octanol). Therefore it is often referred to as the octanol-water partition coefficient. An extension of the Log P is the distribution coefficient Log D which also takes into account the ionic species of a specific compound at a defined pH. Therefore the Log D is preferably used as an indicator of the solubility of a compound in water when the compound of interest is known to be ionized at a certain pH. Hence, K_{OC} is related to the Log D (Xing & Glen, 2002).

The Log D is calculated as follows:

For acids

$$\log D = \log P - \log[1 + 10(pH - pKa)]$$

For bases

$$\log D = \log P - \log[1 + 10(pKa - pH)]$$

where log P is the partition coefficient, and pKa is the ionization constant (Xing & Glen, 2002).

A value of Log D < 0 indicates that the compound is highly polar (hydrophilic) and therefore well soluble in water. If the Log D > 3, the compound is hydrophobic, which means it is not well soluble in water and prone to adsorption (Sangster, 1997). **Table 2** summarizes the properties of selected pesticides.

Common practices to remove pesticides from both drinking water and wastewater, together with other trace organic compounds (TROCs) such as pharmaceuticals, personal care products and industrial chemicals, are adsorption on activated carbon (Kennedy *et al*., 2015; Mailler *et al*., 2015), oxidation using ozone, chlorine, H_2O_2 or UV and combination of adsorption and oxidation (Broséus *et al*., 2009; Derco *et al*., 2015). Moreover dense membrane filtration systems such as reverse osmosis (RO) and nanofiltration (NF) have been proven to remove pesticides effectively (Bonné *et al*., 2000; Plakas *et al*., 2006). One laboratory scale study in 2014 has proven that membrane distillation (MD) provides a barrier for TROCs (Wijekoon *et al*., 2014). In **Table 3** the advantages and disadvantages of different technologies used for removal of pesticides from drinking water are listed and evaluated.

Table 2: Physicochemical properties of the selected pesticides

	Use	Log D at pH 7	Vapour pressure at 25ºC [mmHg]	Water solubility at pH 7, 25ºC [mg/L]	K_{OC} pH 7	Polarity at pH 7	Molecular weight [g/Mol]	pK_H 25 °C
Phorate	Insecticide	3.67	2.60×10^{-3}	50	2,360	Apolar	260.38	4.75
Parathionmethyl	Insecticide	2.82	2.4×10^{-4}	37	816	Apolar	263.21	5.65
Atrazine	Herbicide	2.64	1.27×10^{-5}	69	647	Apolar	215.68	7.28
Dichlorvos	Insecticide	1.07	1.45	57,000	91.5	Apolar	220.98	5.13
Clofibric acid	Herbicide	-1.06	1.03×10^{-4}	100,000	1	Polar	214.65	9.54

Table 3: *Technologies for removal of pesticides in drinking water*

Technology	Removal efficiency	Advantage	Disadvantage	Reference
Adsorption on activated carbon (GAC/PAC)	Highly depends on process conditions and water matrix (DOC), as well as selected carbon. e.g atrazine: 50-80% removal with 5 mg/L PAC and 4-5 h contact time, 10% breakthrough at 20,000 to 50,000 BV at EBCT = 7.6 minutes	Relatively high removal efficiency for many pesticides	Selective removal depending on physico-chemical properties such as charge of compound or its molecular mass, competitive adsorption, removal of PAC after process, saturation of activated carbon, non-destructive method	Ormad *et al*., 2008; Snyder *et al*., 2007
Oxidation by chlorine	Highly depends on dosage and water matrix (NOM) 10-50% for triazines. For other pesticides, ranges from 30 to 100% e.g atrazine: <20% removal with 3 mg/L free chlorine and 24 h contact time	Simple application, long term disinfection with regards to bacteria	Very selective removal, formation of trihalomethanes, formation of oxidation by-products, Low oxidizing character (E^0 = 1.36V)	Ormad *et al*., 2008; Snyder *et al*. 2007
Ozonation	50% for triazines, 80% for organophosphorous pesticides	High oxidizing character (E^0 = 2.8V)	Selective removal, formation of trihalomethanes, formation of oxidation by-products	Ormad *et al*., 2008
Chemical precipitation With aluminium sulphate or ferric chloride	e.g. atrazine <20% removal with aluminium sulphate or ferric chloride; DDT 20% removal with aluminium sulphate or ferric chloride		Low removal rate, sludge generation	Ormad *et al.,* 2008; Snyder *et al*., 2007
UV and advanced oxidation process (UV/H_2O_2)	Removal efficiency depends on chemical structure (aromatic compounds) e.g. atrazine main removal with UV 60-70% removal with 1000 mJ/cm^2 and 5 mg/L H_2O_2	Effective process for many substances, removes also bacteria, uncomplicated process	Formation of oxidation by-products, footprint of peroxide	Snyder *et al*., 2007
Tight nano-filtration (NF)	Removal through size exclusion, electrostatic interactions and adsorption on membrane. Therefore, removal efficiency mainly depends on molecular size, polarity, hydrophobicity/ hydrophilicity, and molecular weight cut off of the selected membrane. e.g. atrazine: 50-80% removal	Successful large scale application for pesticide removal in WTP, Méry-sur-Oise, Paris, France	Not advisable to use NF solely for pesticide removal, Hydrophobic substances are not well retained, aging of membranes leads to reduced rejection	Cyna *et al*., 2002; Plakas & Karabelas, 2012; Plakas *et al*., 2006

Table 3 (cont.): Technologies for removal of pesticides in drinking water

Technology	Removal efficiency	Advantage	Disadvantage	Reference
Reverse osmosis (RO)	Removal through size exclusion and electrostatic Interactions, >90% removal for most TROCs including pesticides	Successful large scale application for pesticide removal in WTP, Amsterdam, Leiduin Netherlands	Energy demand, pre-treatment of water, remineralisation	Bonné *et al*., 2000; Snyder *et al*., 2007
Membrane distillation (MD)	Strongly depends on vapour pressure and Log D, e.g. atrazine >95% removal	Very high removal rate of non-volatile compounds	Energy demand, remineralisation, only proven in laboratory scale	Wijekoon *et al*., 2014

References

Arias-Estévez, M., López-Periago, E., Martínez-Carballo, E., Simal-Gándara, J., Mejuto, J.-C. & García-Río, L. (2008). The mobility and degradation of pesticides in soils and the pollution of groundwater resources. *Agriculture, Ecosystems & Environment, 123*(4), 247-260.

Ayuso-Gabella, N., Page, D., Masciopinto, C., Aharoni, A., Salgot, M. & Wintgens, T. (2011). Quantifying the effect of Managed Aquifer Recharge on the microbiological human health risks of irrigating crops with recycled water. *Agricultural Water Management, 99*(1), 93-102.

Bennett, A. (2008). Drinking water: Pathogen removal from water–technologies and techniques. *Filtration & Separation, 45*(10), 14-16.

Bonné, P.A.C., Beerendonk, E.F., van der Hoek, J.P. & Hofman, J.A.M.H. (2000). Retention of herbicides and pesticides in relation to aging of RO membranes. *Desalination, 132*(1-3), 189-193.

Broséus, R., Vincent, S., Aboulfadl, K., Daneshvar, A., Sauvé, S., Barbeau, B. & Prévost, M. (2009). Ozone oxidation of pharmaceuticals, endocrine disruptors and pesticides during drinking water treatment. *Water Research, 43*(18), 4707-4717.

Central Pollution Control Board. (2008). *Status of groundwater quality in India (Part-II).* Delhi: Central Pollution Control Board, Ministry of Environment and Forests, Government of India.

Chakraborti, D., Das, B. & Murrill, M.T. (2011). Examining India's Groundwater Quality Management. *Environmental Science & Technology, 45*(1), 27-33.

Chaudhary, V., Jacks, G. & Gustafsson, J-E. (2002). An analysis of groundwater vulnerability and water policy reform in India. *Environmental Management and Health, 13*(2), 175-193.

Cyna, B., Chagneau, G., Bablon, G. & Tanghe, N. (2002). Two years of nanofiltration at the Méry-sur-Oise plant, France. *Desalination, 147*(1-3), 69-75.

Derco, J., Dudáš, J., Valičková, M., Šimovičová, K. & Kecskés, J. (2015). Removal of micropollutants by ozone based processes. *Chemical Engineering and Processing-Process Intensification, 94*, 78-84.

Kennedy, A.M., Reinert, A.M., Knappe, D.R.U., Ferrer, I. & Summers, R.S. (2015). Full- and pilot-scale GAC adsorption of organic micropollutants. *Water Research, 68*, 238-248.

Lari, S.Z., Khan, N.A., Gandhi, K.N., Meshram, T.S. & Thacker, N.P. (2014). Comparison of pesticide residues in surface water and ground water of agriculture intensive areas. *Journal of Environmental Health Science & Engineering, 12*, 11. doi:10.1186/2052-336X-12-11

Levantesi, C., Bonadonna, L., Briancesco, R., Grohmann, E., Toze, S. & Tandoi, V. (2012). *Salmonella* in surface and drinking water: Occurrence and water-mediated transmission. *Food Research International, 45*(2), 587-602.

Mailler, R., Gasperi, J., Coquet, Y., Deshayes, S., Zedek, S., Cren-Olive, C., Cartiser, N., Eudes, V., Bressy, A., Caupos, E., Moilleron, R., Chebbo, G. & Rocher, V. (2015). Study of a large scale powdered activated carbon pilot: Removals of a wide range of emerging and priority micropollutants from wastewater treatment plant effluents. *Water Research, 72*, 315-330.

McFarlane, D.J. & Williamson, D.R. (2002). An overview of water logging and salinity in southwestern Australia as related to the 'Ucarro' experimental catchment. *Agricultural Water Management, 53*(1-3), 5-29.

Mnif, W., Hassine, A.I.H., Bouaziz, A., Bartegi, A., Thomas, O. & Roig, B. (2011). Effect of Endocrine Disruptor Pesticides: A review. *International Journal of Environmental Research and Public Health, 8*(6), 2265-2303.

Mutiyar, P.K., Mittal, A.K. & Pekdeger, A. (2011). Status of organochlorine pesticides in the drinking water well-field located in the Delhi region of the flood plains of river Yamuna. *Drinking Water Engineering and Science, 4*, 51-60.

Ormad, M.P., Miguel, N., Claver, A., Matesanz, J.M. & Ovelleiro, J.L. (2008). Pesticides removal in the process of drinking water production. *Chemosphere, 71*(1), 97-106.

Plakas, K.V. & Karabelas, A.J. (2012). Removal of pesticides from water by NF and RO membranes - A review. *Desalination, 287*, 255-260.

Plakas, K.V., Karabelas, A.J., Wintgens, T. & Melin, T. (2006). A study of selected herbicides retention by nanofiltration membranes - The role of organic fouling. *Journal of Membrane Science, 284*(1-2), 291-300.

Sangster, J. (1997). *Octanol-Water Partition Coefficients: Fundamentals and Phyical Chemistry.* New York: John Wiley & Sons .

Sankararamakrishnan, N., Sharma, A.K. & Sanghi, R. (2005). Organochlorine and organophosphorous pesticide residues in ground water and surface waters of Kanpur, Uttar Pradesh, India. *Environment International, 31*(1), 113-120.

Sikarwar, S., Gandhi, K., Singh, K., Kashyap, S.M., Khan, S. & Thacker, N. (2014). Assessment of pesticide contamination in ground water from intensive agricultural sites, using Gas Chromatography-Mass Spectrometry. *Journal of Environmental Science and Engineering, 56*(2), 161-168.

Singh, K.P., Malik, A., Mohan, D. & Takroo, R. (2005). Distribution of persistent organochlorine pesticide residues in Gomti river, India. *Bulletin of Environmental Contamination and Toxicology, 74*(1), 146-154.

Snyder, S.A., Adham, S., Redding, A.M., Cannon, F.S., DeCarolis, J., Oppenheimer, J., Wert, E.C. & Yoon, Y. (2007). Role of membranes and activated carbon in the removal of endocrine disruptors and pharmaceuticals. *Desalination, 202*(1-3), 156-181.

Tielemans, M. (2007). Artificial recharge of groundwater in The Netherlands. *Water Practice & Technology, 2*(3). Retrieved from https://doi.org/10.2166/wpt.2007.064

Vermont Department of Health. (2015). Giardia [Online]. Retrieved from http://healthvermont.gov/enviro/water/coliform.aspx

Wijekoon, K.C., Hai, F.I., Kang, J., Price, W.E., Cath, T.Y. & Nghiem, L.D. (2014). Rejection and fate of trace organic compounds (TrOCs) during membrane distillation. *Journal of Membrane Science, 453*, 636-642.

Woodie, M. (2014). *Bacteria and viruses commonly found in drinking water: A breakdown of harmful bacteria and viruses and what techniques are used to treat them [online]*. Retrieved from http://www.watertechonline.com/bacteria-and-viruses-commonly-found-in-drinking-water/

Xing, L. & Glen, R.C. (2002). Novel methods for the prediction of logP, pKa, and logD. *Journal of Chemical Information and Computer Sciences, 42*(4), 796-805.

Yadav, I.C., Devi, N.L., Syed, J.H., Cheng, Z., Li, J., Zhang, G. & Jones, K.C. (2015). Current status of persistent organic pesticides residues in air, water, and soil, and their possible effect on neighboring countries: A comprehensive review of India. *The Science of the Total Environment, 511*, 123-137.

Research Article

A Nonlinear Cumulative Logit Mixed Effects Model with Decomposition of Time Varying Cumulative Odds into Phases: Applications in Cardiac Surgery

[1]Jeevanantham Rajeswaran, [2]Eugene H. Blackstone

[1]*Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, Ohio, USA.*
[2]*Department of Thoracic and Cardiovascular Surgery and Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, Ohio, USA.*

Received: 5 March, 2018, Accepted: 30 Nov, 2018

Abstract:

Longitudinal ordinal responses are often encountered in many medical studies. Evaluating the temporal trend of the probabilities of the individual categories of the ordinal response and identifying the risk factors that influence these trends are the two important statistical analytical steps that can influence medical decision making. A nonlinear cumulative logit mixed effects model is introduced to identify the temporal decomposition of the time-varying cumulative odds and to estimate the time-related probabilities of individual categories of an ordinal response variable. Further, patient-specific risk factors or determinants, whose influence may or may not change with time, are identified. Marginal likelihood approach using adaptive Gauss-Hermite quadrature for integral approximation and quasi-Newton for optimization were used to estimate the regression parameters, and the approach is easily implemented using readily available software. The application of this model is illustrated using two cardiac surgery longitudinal data: i). Longitudinal assessment of graded heart valve regurgitation following aortic valve repair; and ii). Comparison of pain intensity score between minimally invasive and conventional mitral and/or aortic valve surgical procedures.

Key words: longitudinal ordinal response; cumulative logit link; cumulative logit mixed model; time-varying coefficients; multiphase model.

1. Introduction

Often in medical sciences the response of interest is an ordinal categorical variable such as grade of valve leakage or symptom severity. An important goal of longitudinal statistical analyzes is to evaluate the temporal trend of prevalence of each ordinal category over time after a surgical procedure and to identify patient risk factors that influence these outcomes in the short- or long-term after the procedure. That is, risk factors may or may not depend on time. For example, after aortic valve repair, greater left ventricular hypertrophy may be a risk factor for return of aortic valve regurgitation in the short-term, but not in the long-term; that is, its effect diminishes after the first year. Similarly, use of a certain type of repair procedure for the repair may be a risk factor in the long-term, with an effect

Corresponding author: Dr.Jeevanantham Rajeswaran, Department of Quantitative Health Sciences/ JJN3ClevelandClinic, 9500 Euclid Avenue 4195, USA. e-mail: rajeswj@ccf.org

noticeable after the first year of surgery. In contrast, patients' age may be a risk factor regardless of time period. Hence, the regression coefficients for left ventricular hypertrophy and type of anuloplasty are time-varying. Many have studied time-varying coefficient models through a non-parametric approach (Hoover, *et al.,* 1998; Wu & Yu, 2002; Hastie & Tibshirani, 1993) and mainly involved continuous longitudinal data. In this paper, we propose a parametric approach to fit a longitudinal ordinal data.

When one encounters responses that are measured in a series of ordered categories, these ordinal responses requires the use of statistical models that take into consideration of ordinal nature of the data, such as the *cumulative or graded response model* (McCullagh, 1980; McCullagh & Nelder, 1989). When the link function is of logit form, the model is called a *cumulative logistic model*. Since the link function of an ordinal response is already nonlinear, most of the regression modeling of ordinal response focuses on the linear predictors. That is, regression coefficients enter the model linearly in the cumulative logit domain. A detailed overview of analysis of ordinal categorical data is given in, for example, Liu & Agresti (2005)

When longitudinal (repeated) measurements are available for each patient, the two most commonly used longitudinal statistical approaches are:

1) marginal (population-average) models, in which the repeated nature of the data within each patient is modeled by a joint dependence structure, and
2) mixed effects models (also known as conditional, subject-specific, or multilevel), in which, in the simplest version, a random effect is introduced for each patient (Laird & Ware, 1982; Diggle *et al.,* 2002).

These random effects represent the influence of a patient on his or her longitudinal observations that is not captured by the observed covariates. Further, in observational clinical outcomes studies, the longitudinal data is often highly unbalanced, in the sense that, different number of measurements were observed at different time points for each patient. Many (Fitzmaurice *et al.,* 2011; Lee & Nelder, 2004) have compared these two approaches and discussed interpretation of each. In this paper, since our data is highly unbalanced observational data, we use a mixed effects model approach.

One often overlooked aspect in existing longitudinal models using these two approaches is that the temporal pattern is not explicitly modeled, but is incorporated in only simple ways (such as polynomials, and other simple transformations) within the regression structure. Factors whose influence is time dependent typically require a series of interaction terms with time, which is statistically sub-optimal, especially in observational studies in which there are a large number of covariates to be considered. Rajeswaran & Blackstone (2017) have proposed a multiphase nonlinear mixed effects model for continuous longitudinal data. In this paper, we extend the model for continuous longitudinal data to accommodate ordinal longitudinal data by introducing a nonlinear cumulative logit mixed effects model that enables us to perform patient-specific analysis for longitudinal ordinal data and to estimate the effects of covariates that may or may not depend on time. The main features of this model are :

1) decomposition of cumulative odds into different overlapping time phases (for example, early, constant and late);
2) identification of risk factors influencing each of these phases; and

3) recognition that some concomitant information can be independent of time (proportional odds across time).

The paper is organized as follows. Section 2 briefly reviews a generic function of time that is used in the decomposition model, Section 3 introduces the multiphase cumulative logit mixed effects model, Section 4 briefly discuss data analysis strategy employed, and the model is applied to longitudinal assessment of aortic regurgitation (heart valve leakage) grades after aortic valve repair and to pain score categories after mitral and/or aortic valve surgical procedure.

2. Multiphase time function

We now briefly discuss a family of nonlinear function of time that is being used in our nonlinear mixed effects model in Section 3. The generic equation for the multiphase model was originally described as a model of cumulative mortality by (Hazelrig *et al.,* 1982; Blackstone *et al.,* 1986). The main motivation to use this generic time function to model the cumulative odds is that, often, in time to event analyzes, particularly, in intervention studies such as cardiac surgery, the hazard of adverse events is high right after the surgery, then the hazard gradually decreases and remained low and stable during an intermediate period of time, and then it starts to increase again. Blackstone *et al.,* (1986) showed that these type of time-varying hazard is easily modeled using the time function described below. In our longitudinal data analysis experience using cardiac surgery data, we also found this type of temporal behavior in the longitudinal outcomes. Further, this is a very flexible nonlinear function that can handle almost any shapes, including linear functions. We translate these equations into the cumulative odds domain in our model. The family of equations is given as,

$$G(t,\boldsymbol{\psi}) = \frac{|\lambda|-\lambda}{2|\lambda|} + \frac{\lambda}{|\lambda|} + \left[1 + \phi(\alpha)\left(\frac{|\alpha|-\alpha}{2|\alpha|} + \frac{|\lambda|t}{\rho(t_{1/2})}\right)^{-\frac{1}{\lambda}}\right]^{-\frac{1}{\alpha}} \quad (1)$$

where $\alpha > 0$ and/or $\lambda > 0$, $\phi(\alpha) = \alpha$ if $\alpha > 0$, and $\phi(\alpha) = -1$ if $\alpha < 0$. Shaping parameter $\boldsymbol{\psi} \equiv (\alpha, \lambda, t_{1/2})$, with $G(t_{1/2}) = 1/2$ and $t_{1/2} > 0$. Natural constraints on G are that $G(0, \boldsymbol{\psi}) = 0$ and $G(t, \boldsymbol{\psi}) \to 1$ as $t \to \infty$. $G(t, \boldsymbol{\psi}) > 0$ for $\alpha > 0$ and/or $\lambda > 0$. However, when $\alpha < 0$ and $\lambda < 0$, $G(0) \neq 0$. Hence, $G(t, \boldsymbol{\psi})$ does not exists for $\alpha < 0$ and $\lambda < 0$. Thus, when $\alpha > 0$ and $\lambda > 0$; or $\alpha > 0$ and $\lambda < 0$; or $\alpha < 0$ and $\lambda > 0$ the function (1) reduces to simpler forms. The function can be further simplified when $\alpha \to 0^+$ or $\lambda \to 0^+$. These simplifications play important role in parameter estimation, when α or λ becomes very small. See Rajeswaran & Blackstone, (2017) for details of these function and their limiting behaviors. The first derivative of G with respect to t, $a(t, \boldsymbol{\psi}) = \partial G/\partial t$ is non-negative. Thus, G is non-decreasing. $G(t, \boldsymbol{\psi})$ or $a(t, \boldsymbol{\psi})$ or transformations of these can be used as a time function in our multiphase model. In our experience in analyzing a number of longitudinal data from cardiac surgery, the most commonly used functions are $a(t, \boldsymbol{\psi})$ for the early phase and

$$b(t,\boldsymbol{\psi}) = \frac{a(t,\boldsymbol{\psi})}{1-G(t,\boldsymbol{\psi})} \quad (2)$$

for the late phase. Figure 1 shows some typical forms of cumulative odds for these two functions.

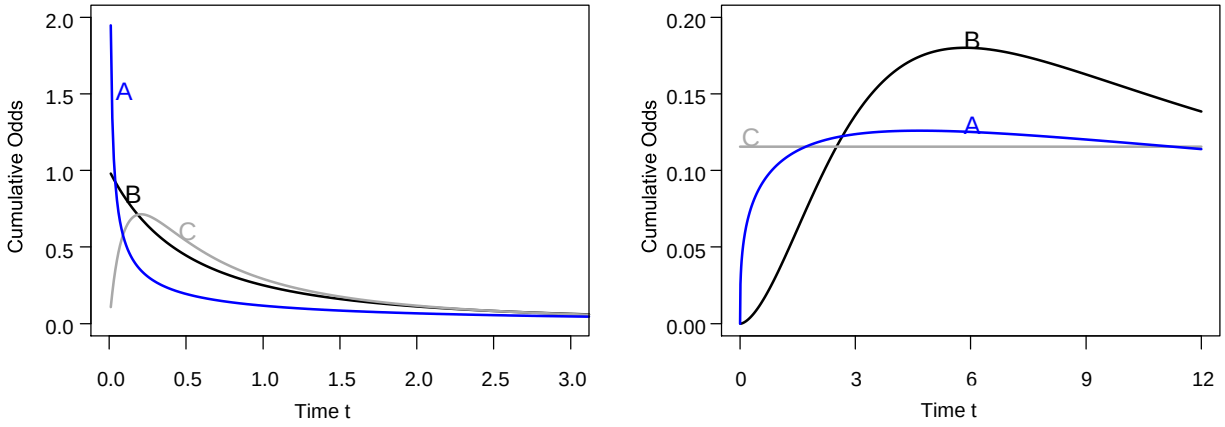

Figure 1 : *Typical forms of cumulative odds. Left plot depicts the function a(t, $\boldsymbol{\psi}$), with $t_{1/2} = 1$. (A) starts at an infinite value: $\alpha = 0.5, \lambda = -2$; (B) early peaking: $\alpha = 0.5, \lambda = 1$; (C) starts at a finite value: $\alpha = 1, \lambda = 1$. Right plot depicts b(t, $\boldsymbol{\psi}$), with $t_{1/2} = 6$. (A) late plateauing: $\alpha = -0.75, \lambda = 0.2$; (B) late peaking: $\alpha = 0.75, \lambda = 0.5$; (C) constant: $\alpha = -1, \lambda = 0$.*

3. A nonlinear cumulative logit mixed effects model

Assume a C-category ordinal response **Y** with Y_{ij} be the j^{th} $(j = 1,\ldots,k_i)$ ordinal longitudinal response for subject i $(i=1,\ldots,n)$, then conditional cumulative probabilities for the C categories of outcome Y_{ij} are denoted as $\pi_{ijc} = P(Y_{ij} \leq c|U_i, \boldsymbol{x}_{ij}, t_{ij}) = \sum_{h=1}^{c} P_{ijh}$, where U_i is the random intercept, $\boldsymbol{x}_{ij}$ is the concomitant information, t_{ij} is the time point at which response Y_{ij} observed, and P_{ijh} represents the conditional probability of response in category h. Suppose, there are L overlapping time phases of temporal trend of cumulative odds, we propose the following generalized nonlinear cumulative logit mixed model for conditional cumulative probabilities:

$$(3) \qquad logit(\pi_{ijc}) = log\left(\frac{\pi_{ijc}}{1-\pi_{ijc}}\right) = \gamma_c + \boldsymbol{X}_0\boldsymbol{\beta}_0 + log\left(\sum_{l=1}^{L} \mu_l(\boldsymbol{X}_l, \boldsymbol{\beta}_l)\, T_l(\boldsymbol{t}, \boldsymbol{\psi})\right) + U_i,$$

$$c = 1, \ldots . C - 1$$

where the intercepts $\{\gamma_c\}$ are often called *cutpoint* parameters. They are non-decreasing in c, because the cumulative logit is an increasing function of π_c, which itself is increasing in c for fixed concomitant information $\boldsymbol{X}$ and time $\boldsymbol{t}$. This model assumes that the log-odds ratio of the cumulative probability is the same for all categories, that is, identical effect of the predictors for each cumulative probability. McCullagh, (1980) called this assumption of identical odds ratios across $C - 1$ cutoffs the *proportional odds* assumption. That is, this nonlinear proportional odds mixed effects model expresses the ordinal responses in C categories in terms of $C - 1$ cumulative logits. See for example, Rajeswaran & Blackstone (2007) for description of cumulative logit link models.

Xs are matrices of covariates that are not necessarily equivalent, and **β**s are their fixed-effect parameters. The U_is are independent patient-specific random intercepts normally distributed as $N(0, \boldsymbol{\sigma_u^2})$, and, given the random intercept U_i, the conditional distribution of Y_{ij} is multinomial. That is, we assume conditional independence. $T_l(t, \boldsymbol{\psi})$ is a function of time that depends only on time t and shaping parameter **ψ**, and this function can be any of the forms or transformations of G(t, **ψ**) (1) in Section 2. The model (3) has two main components of fixed covariates: the first is an overall mixed effects model (namely, $\boldsymbol{X_o\beta_o} + U_i$) that does not depend

on time, and the second is a series of log-linear models $\mu_l(\boldsymbol{X_l}, \boldsymbol{\beta_l}) = exp\{\boldsymbol{X_l\beta_l}\}$ $(l = 1,...,L)$ that are modulated by time function $T_l(t, \boldsymbol{\psi})$. Hence, the effects of covariates $\mathbf{X}_l$s are time varying. In the cumulative odds domain, the model (3) can be written as

$$\frac{\pi_{ijc}}{1-\pi_{ijc}} = \exp(\gamma_c + \boldsymbol{X}_0\boldsymbol{\beta}_0 + U_i) \times \sum_{l=1}^{L} \mu_l(\boldsymbol{X}_l, \boldsymbol{\beta}_l)\, T_l(\boldsymbol{t}, \boldsymbol{\psi}) \quad (4)$$

Hence, multiple overlapping L phases of risk are additive in the cumulative odds domain, with each phase individually shaped by a function of time $T_l(t, \boldsymbol{\psi})$ and scaled by a function of concomitant information $\mu_l(\boldsymbol{X_l}, \boldsymbol{\beta_l})$. Note that we can use as many phases as the data warrant, but in our data analysis experience, 3 phases – early, constant and late – are usually adequate. We now briefly describe these 3 commonly occurring phases.

Early Phase: The time function $a(t, \boldsymbol{\psi})$ is the most commonly used as function $T_l(t, \boldsymbol{\psi})$ in this phase. Under this scenario, the scaling parametric function $\mu_l(\mathbf{X}_l, \boldsymbol{\beta}_l)$ is related to the area beneath function $a(t, \boldsymbol{\psi})$.

Constant Phase: This phase is time independent. Hence, it will only have concomitant information.

Late Phase: The time function $b(t, \boldsymbol{\psi})$ is the most commonly used as function $T_l(t, \boldsymbol{\psi})$in this phase. One of the interesting properties of $b(t, \boldsymbol{\psi})$ or any of its transformation is that the parameter $t_{1/2}$ behaves like a scalar. Hence, we introduce the concomitant information through $t_{1/2}$ as follows: $t_{1/2} = \mu_l(\mathbf{X}_l, \boldsymbol{\beta}_l)$.

Note that for terminology convenience, we have identified $a(t, \boldsymbol{\psi})$ with early phase and $b(t, \boldsymbol{\psi})$ with late phase. As we will see in Example 2, however, this is not always the case. Any equation can be used at any phase of the time as data warrant.

3.1 Estimation

We use the method of marginal likelihood (Diggle *et al.,* 2002) to estimate the shaping parameters and concomitant information coefficients and the corresponding variances. Let $\boldsymbol{\beta} = (\beta_0, \beta_1,...,\beta_L)$ be a vector of regression parameters; $\boldsymbol{\psi} = (\psi_1,...,\psi_L)$ be a vectors of shaping parameters; and let $\boldsymbol{\Theta}=(\boldsymbol{\beta}, \boldsymbol{\psi}, \boldsymbol{\sigma_u^2})$. Further, we know that $P_{ijc} = P(Y_{ij} \le c | U_i, \boldsymbol{x}_{ij}, t_{ij}) - P(Y_{ij} \le c-1 \,|\, U_i, \boldsymbol{x}_{ij}, t_{ij})$ be the conditional probability of individual category c, $(c = 1,...,C)$, for the ordinal longitudinal response for subject i, $(i = 1,...,n)$, observed at time t_{ij}, $(j = 1,...,k_i)$. Then the marginal likelihood is given by,

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\Theta}|\mathbf{y}_i) &= \prod_{i=1}^{n} \int_{\mathcal{R}} f(\mathbf{y}_i, U_i|\boldsymbol{\Theta}) dU_i \\
&= \prod_{i=1}^{n} \int_{\mathcal{R}} \prod_{j=1}^{k_i} f_{y|U_i}(y_{ij}|U_i; \boldsymbol{\beta}, \boldsymbol{\Psi}) f_U(U_i; \sigma_u^2) dU_i \\
&= \prod_{i=1}^{n} \int_{\mathcal{R}} \prod_{j=1}^{k_i} \prod_{c=1}^{C} P_{ijc}^{I(y_{ij}=c)} (2\pi\sigma_u^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_u^2} U_i^2\right\} dU_i,
\end{aligned}$$

where $\mathbf{y}_i = (y_{i1},...,y_{iki})^{\mathrm{T}}$ is the vector of observed responses for subject i, $(i = 1,...,n)$, $f_{y|Ui}$ is the conditional density of the longitudinal response, f_u is the density of the random effects and $I(y_{ij} = c) = 1$ if $y_{ij} = c$ and 0 otherwise.

In other words, simply speaking. Because random effects are unknown, we obtain the marginal likelihood function of **Y**, merely by integrating the joint distribution of **Y** and U with respect to U. That is, we construct the usual product of multinomials that would apply if U were known, and then integrate out the random effect U. Because the integral does not have a closed form, we use an approximation for the likelihood and maximize the approximated likelihood using standard numerical methods, for example, quasi-Newton.

This can be implemented by using PROC NLMIXED (SAS®, Inc., Cary, NC). This procedure uses adaptive Gauss-Hermite quadrature (Pinheiro & Bates, 1995) for numerical integration with respect to the random effect to determine the marginal likelihood function. Although PROC NLMIXED is not naturally designed for multinomial responses, we can use it for ordinal models by specifying the form of the likelihood, which is allowed in this SAS procedure. As a note of caution when using this procedure, a) care should be given to numerical integration when the dimension of U is greater than one when analyzing non-normal longitudinal responses, as in this paper; b) the only distribution available for the random effects is normal. SAS codes for example 2 is given in Appendix A.

Note that, by using (3), we can estimate the patient-specific conditional probabilities $\hat{P}_{ijh}$ for the individual category h. Further, the estimated marginal probabilities, $\widehat{M}_{jh}$ can be obtained by integrating the conditional probability over the distribution of random intercept U (Fitzmaurice *et al.,* 2011). Since this is a one dimensional integration, we can use *integrate*( ) function in R (R Core Team, 2013) to implement the integration. An approximation to marginal probabilities can also be obtained by averaging the patient-specific conditional probabilities, $\widehat{M}_{jh} \approx \frac{1}{n}\sum_{i=1}^{n} \hat{P}_{ijh}$.

4. Data analysis

As we stated in our objectives in Section 1, we employed the following two-step strategy to analyze data.

Step 1 *Temporal Trend*:
We explicitly model the temporal pattern without considering any covariate information in (3). That is, we consider a model with time t and the longitudinal outcome, and estimate intercepts and the shaping, scaling parameters of the time function and identify the phases.

Remark: We have used the following strategy to identify the phases and to estimate the shaping and scaling parameters: We have started with two phases, with early phase time function $T_1(t, \boldsymbol{\psi}) = a(t, \boldsymbol{\psi})$ with a small starting value for $t_{1/2}$ and late phase time function $T_2(t, \boldsymbol{\psi}) = b(t, \boldsymbol{\psi})$ with a larger starting value for $t_{1/2}$. We have used (1,1), (-1,1) and (1,-1) as starting values for α and λ, respectively, for both phases and noted the convergence, likelihood estimates and parameter estimations. It turns out, for the data analysis in Example 1 given below, $T_1(t, \boldsymbol{\psi})$ function with very small λ, almost 0, which was not statistically significantly different from 0. Hence, we fixed the λ at 0; for $T_2(t, \boldsymbol{\psi})$ function, we observed a very small α, almost 0 and λ almost equals to 1, which was not statistically significantly different from 1. Hence, we fixed the α at 0 and λ at 1, for this phase. See ***Table 1*** for the estimated and fixed values.

Step 2 *Multivariate Analysis*:
Since most of the observational studies contain large number of covariates, more than 50 in some studies, and PROC NLMIXED does not have any variable selection capabilities, we have used the following ad-hoc method to screen the variables. Suppose that, for example, Step 1 yielded two phases with an early phase influencing the trend approximately up to 2

years and a late phase thereafter; To identify the possible candidate for risk factors for each phase, we divided the data (longitudinal response and the corresponding baseline covariate information) into two parts; 0-2 years and greater than 2 years and for each part, we then use cumulative logistic regression and PROC LOGISTICS (stepwise selection with liberal entry criteria of 0.12 and stay criteria of 0.1) to screen the variables. We use all the data and the same liberal criteria to select possible variables for the overall model. Once we have identified possible of candidates for risk factors for the overall model and for the models of the each of the phases, these variables and their transformations, if any, were entered into the models one by one and then eliminated one by one until all the variables remaining had a P value of 0.05 or less.

We now discuss data analysis results from two examples aroused in the field of cardiac surgery.

4.1 Example 1

Aortic valve replacement is the main choice of procedure for patients with severe aortic stenosis. However, appropriate surgical strategy for patients with severe aortic regurgitation (AR, aortic valve leakage) is still debated. Aortic valve regurgitation is frequent in patients with long-standing left-sided valve diseases. Aortic valve repair is a surgical option for adult patients with AR. However, its durability is questionable. It has been suggested that repair should be employed only in carefully selected patients. To assess durability of repair, postoperative AR grades were longitudinally collected from 1,163 echocardiograms in 521 patients who underwent aortic valve repair at Cleveland Clinic from 1984 to 2004 are considered in this retrospective cohort study. AR was graded as none (0), mild (1+), moderate (2+), moderately severe (3+) or severe (4+). Hence, this can be regarded as longitudinal ordinal response data with 5 individual categories. The main objective of this data analysis is to quantify the temporal trend of prevalence of each grade over time after the repair and to identify patient-specific risk factors associated with higher likelihood of having higher grade of postoperative AR.

Overall Temporal Trend

Temporal trend analysis using model (3) yielded 2 phases: an early decreasing phase dominating almost the first 2 years and a late phase that increasing initially and plateaued thereafter. (**Figure 2**). **Figure 2** shows the temporal decomposition of aortic regurgitation grade 0 for a "typical" patient, where the random effect is fixed at 0. The estimates and the standard error of the scaling and shaping parameters and standard deviation of the distribution of *U* are given in the *Table 1*.

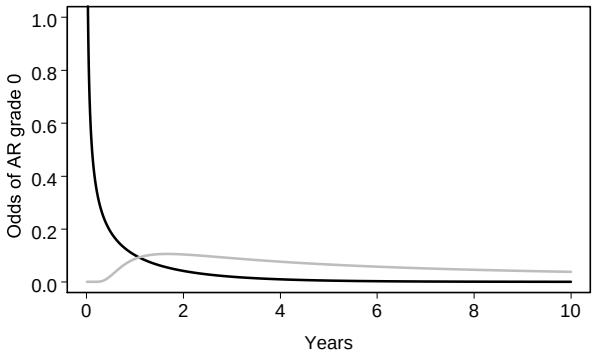

Figure 2 : *Decomposition of odds of aortic regurgitation grade 0 for a "typical" patient* ($U_i = 0$).

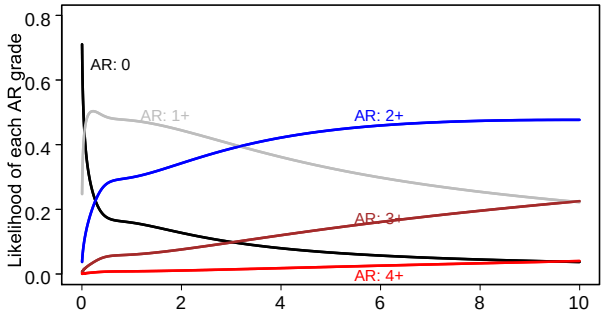

Figure 3 : *Temporal trend of likelihood of each aortic regurgitation regurgitation grade after the repair for a "typical" patient (Ui = 0).*

The overall temporal trend shows that the likelihood (conditional probabilities) of a "typical" patient being in different AR grades changed over time (**Figure 3**). While the likelihood of the patient in grade 0 decreased rapidly within the first six months after the repair and gradually decreased thereafter, likelihood of being in grades 2+ or 3+ increased rapidly during the during the first six months and increased gradually thereafter. Likelihood of a "typical" patient in grade 4+ increased mildly during the 10 years period after AV repair.

Table 1: Estimates of scaling and shaping parameters of the temporal decomposition model. Note that, $\lambda = 0$ means limiting case of $a(t, \psi)$ when $\lambda \to 0^+$ and $\alpha = 0$ means limiting case of $b(t, \psi)$ when $\alpha \to 0^+$. The column "Fixed" identifies the parameters which were fixed to a certain values and which were estimated.

Parameter	Fixed	Estimate ± SE	P
Intercepts (cutpoints)			
γ_1	No	-0.83±0.23	0.0003
γ_2	No	1.4±0.21	<0.0001
γ_3	No	3.4±0.23	<0.0001
γ_4	No	5.6±0.29	<0.0001
Early decreasing phase			
α	No	-2.2±0.45	<0.0001
λ	Yes	0	-
$t_{1/2}$	No	0.38±0.16	0.023
Late increasing phase			
α	Yes	0	-
λ	Yes	1	-
$t_{1/2}$	No	3.8±1.8	0.0025
σ_u^2		5.9±0.85	<0.0001

Multivariate Analysis

Risk factors for higher likelihood of having higher grade of postoperative aortic regurgitation were identified by the parametric model (3). Because of large number of demographic, clinical status, valve pathology, morphology, and etiology variables were considered in this observational study, multivariable analysis was performed as described at the beginning of this section. Data analysis yielded the following results (*Table 2*).

Table 2: Patient-Specific Risk factors for higher likelihood of having higher grade of postoperative aortic regurgitation.

Risk Factor	Coefficient ± SE	P
Overall		
Tricuspid morphology	1.7±0.28	<.0001
Early		
Severe left ventricular dysfunction	0.11±0.039	0.004
Non-coronary cusp prolapse	0.91±0.46	0.05
Constant		
Higher preoperative AR grade	0.77±0.24	0.001

Overall risk factors: Tricuspid AV (compared to Bicuspid AV: **Figure 4**-top left), is appeared to be associated with overall higher likelihood of higher grade of post-op AR; Note that, this is a risk factor that influences the likelihood of post-op higher AR grade, regardless of time;

Early risk factors: Severe left ventricular dysfunction (**Figure 4**) and non-coronary cusp prolapse are appeared to be associated with higher likelihood of having higher grade of early post-op AR. It can be clearly seen from **Figure 4** (top right) that the impact of this factor is early after the surgery;

Late risk factors: Higher preoperative AR grade (**Figure 4**- bottom left), is appeared to be associated with higher likelihood of being in higher grade of late post-op AR. That is, the impact of this variable is late after surgery.

It can be seen from the **Figure 4** that effect of the early-phase risk factor, severe left ventricular dysfunction is larger in the first two years and started to decrease thereafter, on the other hand, the effect of the late-phase risk factor, preoperative AV regurgitation is noticeable after year two.

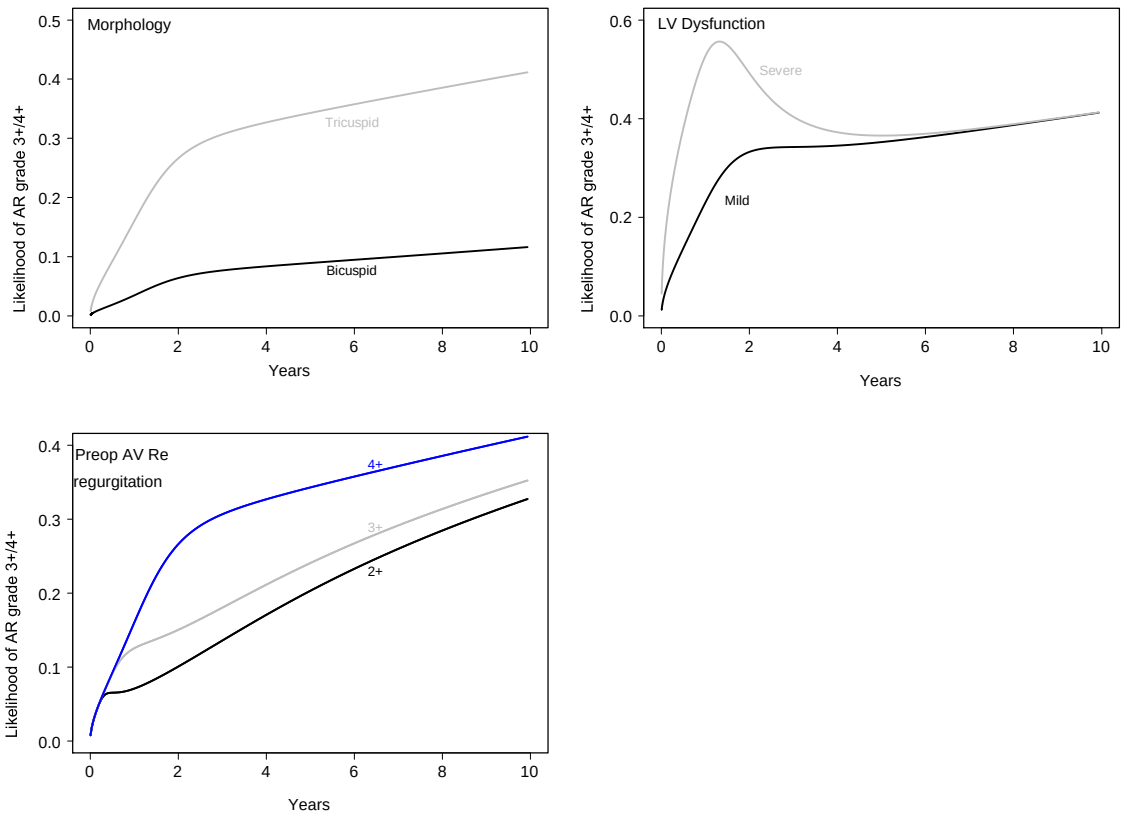

Figure 4: *Likelihood of a "typical" patient being in aortic regurgitation grade 3+ or 4+ after surgery by different risk factors. Top left - influence of morphology: tricuspid vs. biscuspid; top right influence of left ventricular dysfunction; bottom right – influence of preoperative AV regurgitation.*

4.2 Example 2

Full sternotomy for mitral and/or aortic valve operations has been the conventional surgical approach for 50 years; however, in the mid-1990s, minimally invasive approaches were pioneered with the intention of reducing short-term and long-term patient and pain management. To assess the impact of minimally invasive techniques in patients undergoing mitral and/or aortic valve surgical procedures, we performed a propensity-matched comparison of safety and clinical outcomes between patients who had minimally invasive (MIP) and full sternotomy (conventional) approaches to mitral and/or aortic valve surgical procedures. Detail description of the propensity-matched comparison of outcomes in mitral valve, double (aortic and mitral) valve and aortic valve surgical cohorts can be found in Svensson *et al.,* (2010); Atik *et al.,* (2011); and Johnston *et al.,* (2012), respectively. We now briefly describe the propensity score.

The gold standard study design for comparing two treatments is randomized controlled trials (RCTs), where randomization ensures that the treatments groups are balanced with respect to observed and unobserved covariates in the study population. Thus, eliminating selection bias between the treatment groups which, in turns, allows investigators to assess the treatment effect without any bias. However, in observational studies, because of non-randomized assignment/selection of

treatments, distributions of covariates may vary systematically. Hence, to eliminate or reduce the selection bias between the treatment groups before assessing the treatment effect, one has to use some appropriate statistical approaches, such as regression adjustment or stratification, each with their pros and cons. Rosenbaum & Rubin, (1983) in their seminal paper proposed propensity score (PS) strategy to eliminate/reduce selection bias. Simply speaking, PSs are summary of multi-dimensional covariate **X** reduced into a single dimension $P(\mathbf{x})$, propensity (or probability) of receiving a treatment of interest (T). One of the main properties of PS is that PS is a balancing score. This is, $\mathbf{X} \perp T|P(\mathbf{x})$, conditional on PS, treatment T and the distributions of covariates **X** are independent. See, Rosenbaum & Rubin (1983) for further discussion on the properties of PS.

In this paper, we illustrate the application of the model (3) to pain intensity score postoperatively before hospital discharge in the combined mitral and/or aortic valve surgical procedure cohorts. Pain intensity was recorded by nursing staff using the Wong pain-rating scale (Wong & Whaley, 1986) which ranges from none (0) to severe (10). Because of low frequency of high pain scores, scores were collapsed into 5 pain categories: 0 (pain score 0), 1 (pain scores 1-3), 2 (pain scores 4-6), 3 (pain scores 7, 8), and 4 (pain scores 9, 10).

4.2.1 Patients and Propensity-Score Matching

A total of 2797 patients who underwent mitral valve and/or aortic valve using conventional (n=1090 (39%)) or MIP (1707 (61%)) surgical procedure at the Cleveland Clinic between January 2000 to January 2004 are considered in this data analysis. A propensity of having conventional procedure was created using demography, symptoms, valve pathology and etiology, cardiac/non-cardiac comorbidities, coronary anatomy, and procedure variables (44 variables in total). C-Statistic of the propensity model is 0.79, which suggests the model have a good measure of discrimination between the two treatment groups. Using only the propensity scores and greedy matching algorithm (Bergstralh & Konsanke, 1995), conventional patients were matched to MIP patients and identified 783 well matched pairs. **Figure 5**-left shows the distribution of propensity score in conventional and MIP group in the overall and in the matched (shaded area) cohorts. It can be clearly seen that matched pairs covers whole spectrum of propensity score. Covariate balance (D'Agostino, 1998) before and after matching using standardized difference, for some selected variables, is shown in **Figure 5**-right. For example, in the overall cohorts MIP group have more mitral valve degenerative disease patients than the conventional group and difference (bias) is reduced after matching. On the other hand, conventional group is older age than the MIP group patients and difference (bias) is reduced after matching.

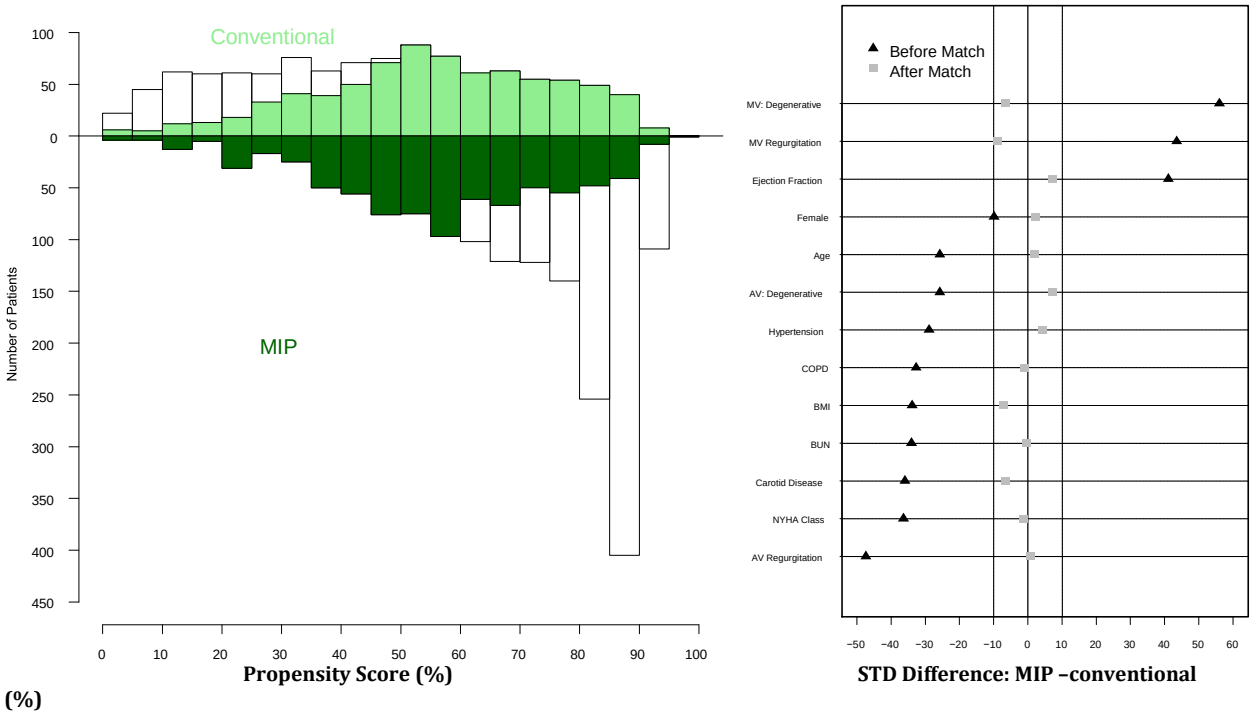

Figure 5: *Left plot depicts mirror histogram of distribution of propensity score in conventional and MIP group with shaded histograms depict the distribution of propensity score in the matched cohort. Right plot depicts covariate balance before and after matching. Key: MV- mitral valve; AV - aortic valve; COPD: chronic obstructive pulmonary disease; BMI - body mass index; BUN - blood uria nitrogen; NYHA class - New York heart association functional class.*

4.2.2 Pain Score: Overall Temporal Trend

In the matched cohort a total of 50029 pain score records collected before discharge are available for the 1566 matched patients. Temporal trend analysis yielded, 3 phases: early, constant, intermediate peaking, and late rising (**Figure 6,** left).

It can be seen from **Figure 6** (right) that percentages of patients (approximate marginal probabilities) in categories 0, 1 (score 2-3), 2 (score 4-6), 3 (score 7-8) and 4 (score 9-10) stayed relatively constant for the first 12 hours. Then, while the percentages of patients in category 0 increased rapidly within the next 4 days, percentages of patients with pain score 1-3, 7-8, and 9-10 decreased slightly and patients in pain score 4-6 decreased rapidly during he same time frame. While percentages of patients with pain score 0 increased gradually, all the other scores decreased slightly thereafter (**Figure *6***, right).

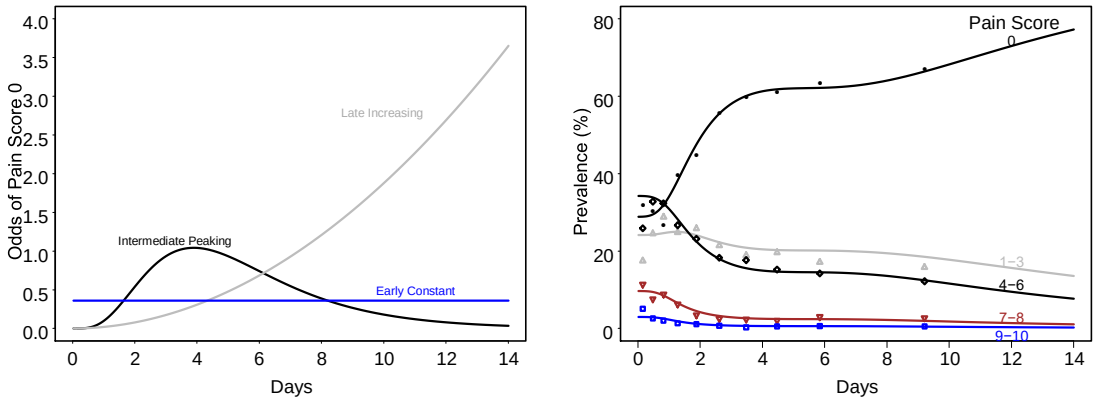

Figure 6 : Left plot depicts decomposition of cumulative odds of pain score category 0. Right plot depicts temporal trend of prevalence of individual pain score categories in the matched cohort. Symbols depict the actual binned averages of each category without regards to the repeated nature of the data, provided here as a crude verification of non-linear model fit.

4.2.3 Pain Score comparison

Pain score comparison between the two approaches in the matched cohort showed that there is an early significant difference (p-value [early constant]<.0001; p-value [intermediate peaking]=0.8; p-value [late rising]=0.6]) in the likelihood of different pain scores between minimally invasive and conventional approaches; with patient who underwent the minimally invasive approach experiencing less early pain than the patients who underwent the conventional approach. That is, in the first 12 hours, minimally invasive approach group appears to be associated with higher percentage of patients with pain score 0 than the conventional approach patients. The percentage of patients with pain score 0 in conventional group appears to increase and become almost equivalent to that of the minimally invasive group after 2 days (**Figure 7**).

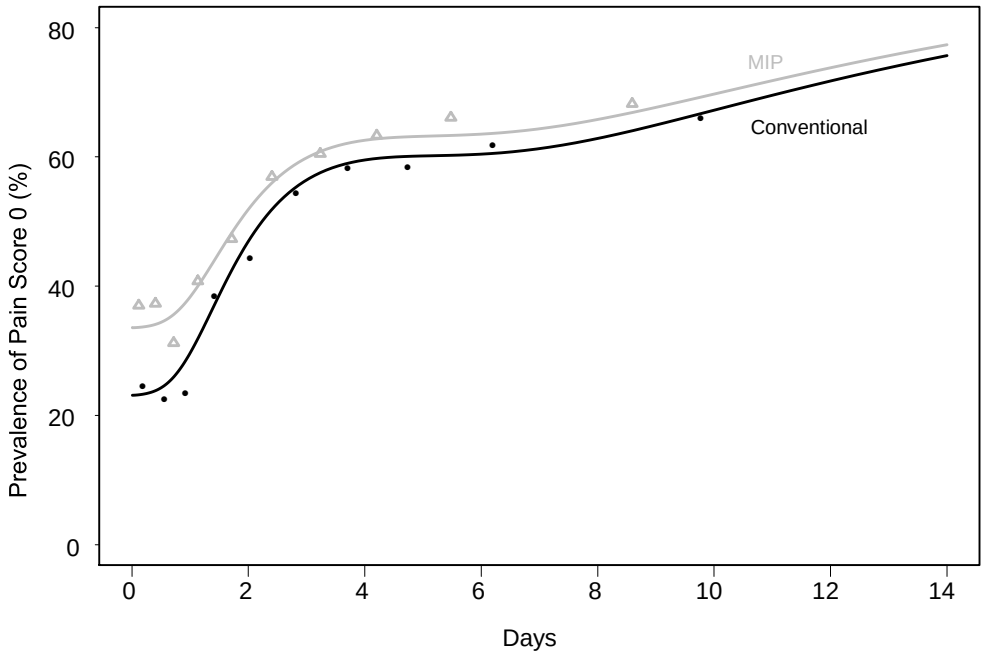

Figure 7: Pain score 0: Comparison between minimally invasive (MIP) and conventional sternotomy groups in the matched cohort. Symbols depict the actual binned averages of pain score 0 without regards to the repeated nature of the data, provided here as a crude verification of non-linear model fit.

5. Summary

This article discusses analysis of ordinal longitudinal data using a non-linear temporal decomposition model. Decomposition of the time-varying pattern of risks into overlapping phases has been developed in the hazard function domain and currently is extensively being used in time-to-event data in cardiac surgery field (Blackstone *et al.,* 1986) and the concept has been extended to model longitudinal continuous data (Rajeswaran & Blackstone, 2017). In this article, we have further extended the concept to longitudinal ordinal data through the cumulative odds domain. In this nonlinear cumulative logit mixed model, we have shown that some factors do not depend on the time (proportional odds assumption across entire period), and some others do (proportional odds only within a time phase, for example, early, constant, or late). Note that one limitation of this model is that currently there is no statistical test available to test the proportional odds assumption. However, as a crude verification, we can use other available models for independent ordinal responses, such as PROC LOGISTIC in SAS, to verify this assumption.

The generic family of equations (Hazelrig *et al.,* 1982) originally used as hazard functions provides a flexible system of mathematical functions for the fitting of temporal pattern. Because our modeling scheme restricts concomitant information to the scaling parameters of each phase, it makes the model mathematically and computationally more flexible.

One of the main advantages of our modeling is that, we have applied our model using a readily accessible standard software. PROC NLMIXED is a readily available, simple-to-use statistical procedure in SAS that takes a conditional (mixed effects) model and estimates parameters through marginal likelihood approach. In addition, it is very easy to embed programming statements to specify complex equations or distributions, such as time function (1). Even though we are using a conditional model approach, we can estimate population average by averaging the conditional estimates.

An advantage over the existing non-parametric time-varying coefficient modeling approaches is that our parametric modeling approach can handle large number of covariates. This is always the case in observational studies. As demonstrated in Example 1, an advantage of formal parametric decomposition of temporal pattern and identifying time-specific risk factors is targeted patient management based on short and long-term risk factors. For example, because the influence of LV dysfunction changes with time and its effect is significant in the first 2 years after surgery, treating physicians can advise such patient to have more periodic follow-ups in the first 2 years after the cardiac surgery. Early phase does not always need to be from equation $a(t, \boldsymbol{\psi})$; in fact, in Example 2, it is a constant. Also noted in Example 2, it is easy to model complex temporal patterns using the generic function or its transformation. The temporal decomposition model also made it simple to detect early differences in pain scores between the surgical groups.

The main limitation of performing multivariate analysis using model (3) is that we have to use an ad-hoc method for variable selection. Variable selection for longitudinal models is a possible future research. Another possible extension of the model (3) is that we can add a random slope to the model or we can introduce a separate random intercepts for each phase model. However, when analyzing a mixed effects model, particularly for non-normal longitudinal data, with more than one dimension of random effects components along with a non-linear time function as in (1), using PROC NLMIXED may sometime produce computational problems and may time consuming. Hence, we need a robust

estimation algorithm to implement the extension and we are working towards this goal.

Acknowledgments

This research was partially supported by a grant from National Institutes of Health - NATIONAL HEART, LUNG, AND BLOOD INSTITUTE, Grant Number: 1R01HL103552

A Appendix: SAS code

In this section we provide SAS code for overall trend (***Figure 6***) in example 2.

```
/*************************************************************************
built        -- Data set
ptid         -- Patient ID
iv_dpain     -- interval to pain score measurement after surgery (in days)
pain_s       -- pain score categories (0, 1, 2, 3, 4)

%decompose   -- SAS macro for the non-linear time function described in Section 2
                when _complt=0 we obtain a( , ) function and when _complt=1 we
                obtain b( , ) function as described in Figure 1
*************************************************************************/
proc sort data=built; by ptid iv_dpain; run;
ods output parameterestimates=peMLu_m;
proc nlmixed data=built gconv=1E-8;
      parms
        /* Cut point parameters */
              a0=1.7119 a1=1.0041 a2=1.8104 a3=1.5433

        /* Early Constant */
              c0=-2.5756

        /* Intermediate peaking */
              thalf1=4.9885 m1=-0.1967

        /* Late Increasing */
              thalf2=8.7578 nu2=-0.3488

        /* Variance of Random effect */
              logsig=-0.08130;

        bounds a1>0, a2>0, a3>0;

        /* Early Constant */
              muc=exp(c0);

        /* Intermediate peaking */
               nu1=0;
               %decompos(_complet=0, _time=iv_dpain, _thalf=thalf1, _nu=nu1, _m=m1);
               g1=_hazard;

        /* Late Increasing */
               m2=0;
               %decompos(_complet=1, _time=iv_dpain, _thalf=thalf2, _nu=nu2, _m=m2);
               g2=_hazard;

        eta=exp(u)*(g1 + muc +g2);
```

```
/* Define the log likelihood by using cumulative odds*/
if (pain_s=0) then do;
     co0=eta*exp(a0);
     z=( co0/(1+co0));
                    end;
else if (pain_s=1) then do;
     co0=eta*exp(a0); co1=eta*exp(a0+a1);
     z=( co1/(1+co1) ) - ( co0/(1+co0) );
                                          end;
else if (pain_s=2) then do;
     co1=eta*exp(a0+a1); co2=eta*exp(a0+a1+a2);
     z=( co2/(1+co2) ) - ( co1/(1+co1) );
                                          end;
else if (pain_s=3) then do;
     co2=eta*exp(a0+a1+a2); co3=eta*exp(a0+a1+a2+a3);
     z=( co3/(1+co3) ) - ( co2/(1+co2) );
                                          end;
else do;
     co3=eta*exp(a0+a1+a2+a3);
     z=1- ( co3/(1+co3) );
                          end;
if (z > 1e-8) then ll=log(z);
else ll=-1e100;

model pain_s ~ general(ll);

random u ~normal(0, exp(2*logsig)) subject=ptid out=randout;

run;
```

References

Atik, F., Svensson, L.G., Blackstone, E.H., Gillinov, A.M., Rajeswaran, J. & Lytle, B.W. (2011). Less invasive versus conventional double valve surgery: A propensity-matched comparison . *Journal of Thorac Cardiovascular Surgery, 141*(6), 1461-1468.

Bergstralh, E.J. & Konsanke, J.L. (1995). *Computerized Matching of Cases to Controls.* Arizona: Mayo Foundation.

Blackstone, E.H., Naftel, D.C. & Turner, M.E. (1986). The Decomposition of TimeVarying Hazard Into Phases Each Incorporating a Separate Stream of Concomitant Information. *Journal of the American Statistical Association, 81*(395), 615-624.

D'Agostino, R.B. (1998). Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine, 17*(19), 2265 - 2281.

Diggle, P.J., Heagerty, P., Liang, K-Y. & Zeger, S.L. (2002). *Analysis of Longitudinal Data* (2nd ed.). Oxford: Oxford University press.

Fitzmaurice, G.M., Laird, N.M. & Ware, J.H. (2011). *Applied Longitudinal Analysis* (2nd ed.). New Jersey: John Wiley & Sons.

Hastie, T. & Tibshirani, R. (1993). Varying-Coefficient Models. *Journal of Royal Statistical Society B, 55*(4), 757-796.

Hazelrig, J.B., Turner, M.E. & Blackstone E. H. (1982). Parametric Survival Analysis Combining Longitudinal and Cross-sectional Censored and Interval-Censored Data with Concomitant information. *Biometrics, 39*(1), 1-15.

Hoover, D.R., Rice, J.A., Wu, C.O. & Yang, L-P. (1998). Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data. *Biometrika, 85*(4), 809-822.

Johnston, D.R., Atik, F.A., Rajeswaran, J., Blackstone, E.H., Nowicki, E.R., Sabik J.F. 3rd, Mihaljevic, T., Gillinov, A.M., Lytle, B.W. & Svensson, L.G. (2012). Outcomes of less invasive J-incision approach to aortic valve surgery. *Journal of Thorac Cardiovascular Surgery, 144*(4), 852-858.

Laird, N.M. & Ware, J.H. (1982). Random-Effects Models for longitudinal data. *Biometrics, 38*(4), 963-974.

Lee, Y. & Nelder, J.A. (2004). Conditional and Marginal Models: Another View. *Statistical Science, 19*(2), 219-238.

Liu, I. & Agresti, A. (2005). The analysis of ordered categorical data: An overview and a survey of recent developments. *Journal of the Spanish Society of Statistics and Operations Research - TEST, 14*(1), 1-73. Retrieved from https://doi.org/10.1007/BF02595397

Matsuyama, K., Matsumoto, M., Sugita, T., Nishizawa, J., Tokuda, Y. & Matsuo, T. (2003). Predictors of residual tricuspid regurgitation after mitral valve surgery. *The Annals of Thoracic Surgery, 75*(6), 1826-1828.

McCarthy, P.M., Bhudia, S.K., Rajeswaran, J., Hoercher, K.J., Lytle, B.W., Cosgrove, D.M. & Blackstone, E.H. (2004). Tricuspid valve repair: durability and risk factors for failure. *Journal of Thorac Cardiovascular Surgery, 127*(3), 674-685.

McCullagh, P. & Nelder, J.A. (1989). *Generalized Linear Models* (2nd ed.). New York: Chapman & Hall/CRC.

McCullagh, P. (1980). Regression Models for Ordinal Data. *Journal of the Royal Statistical Society B, 42*(2), 109-142.

Pinheiro, J.C. & Bates, D.M. (1995). Approximations to the Log-likelihood Function in the Non-linearMixed-Effects Model. *Journal of Computational and Graphical Statistics, 4*(1), 12-35.

R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org/

Rajeswaran J. & Blackstone, E.H. (2017). A Multiphase Non-Linear Mixed Effects Model: An Application to Spirometry after Lung Transplantation. *Statistical Methods in Medical Research, 26*(1), 21-42.

Rajeswaran, J. & Blackstone, E.H. (2007). Interval estimation for individual categories in cumulative logit models. *Statistics in Medicine, 26*(22), 4150-4162.

Rosenbaum, P.R. & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55.

Svensson, L.G., Atik, F., Cosgrove, D.M., Blackstone, E.H., Rajeswaran, J., Krishnaswamy, G., Jin, U., Gillinov, A.M., Griffin, B., Navia, J.L., Mihaljevic, T. & Lytle, B,W. (2010). Minimally invasive versus conventional mitral valve surgery: A propensity-matched comparisonconventional mitral valve surgery: A propensity-matched comparison. *Journal of Thorac Cardiovascular Surgery, 139*(4), 926-932.

Wong, D. & Whaley, L. (1986). Clinical handbook of pediatric nursing. St.Louis:C.V. Mosby Company.

Wu, C.O. & Yu, K.F. (2002). Nonparametric Varying-Coefficient Models for the Analyses of Longitudinal Data. *International Statistical Review, 7*(3), 373-393.

Research Article

Atomistic Model Approach to Identify Defects, Lithium Ion Diffusion and Trivalent Dopants in Li_2MnO_2

[1,2]Navaratnarajah Kuganathan*, [3]Sashikesh Ganeshalingam, [3]Poobalasingam Abiman and [1,2]Alexander Chroneos

[1]*Department of Materials, Imperial College London, United Kingdom.*
[2]*Faculty of Engineering, Environment and Computing, Coventry University, United Kingdom*
[3]*Depratment of Chemistry, University of Jaffna, Sri Lanka.*

Received: 15 March, 2018, Accepted: 16 Dec, 2018

Abstract:

Layered lithium-rich metal oxides have attracted great interest as potential cathode materials for Li ion batteries due to their high Li content required for high energy density. Using atomistic simulation techniques based on classical pair potentials, we calculate intrinsic defects, lithium ion diffusion paths together with activation energies and trivalent doping in Li_2MnO_2. The most favourable intrinsic defect type is found to be the cation anti-site defect, in which Li and Mn ions exchange their positions. Lithium ions diffuse via a zig-zag path with very low activation energy of 0.16 eV. Trivalent dopant Sc^{3+} on Mn site is energetically favourable and could be a synthesis strategy to increase the Li vacancy concentration in Li_2MnO_2.

Keywords: Defects, Diffusion, Dopants, Atomistic simulation

1. Introduction

The interest in solid-state lithium batteries is driven by the increasing requirement for better capacity, cycle performance, safety, and durability. This motivates the research interest on the discovery and application of advanced electrolyte and cathode materials (Tarascon & Armand, 2001; Bruce *et al.,* 2011; Jay *et al.,* 2015; Fisher *et al*., 2013; He *et al*., 2017; Armstrong *et al*., 2011; Kuganathan & Islam, 2009; Islam & Fisher, 2014; Kordatos *et al*., 2018; Kuganathan *et al.,* 2018a; Kuganathan *et al.,* 2018b; Kuganathan *et al.,* 2018c; Kuganathan *et al.,* 2018d; Kuganathan *et al*., 2018e).

A key area is the identification and study of alternative cathode materials for rechargeable Li-ion batteries. These cathode materials are required to have high energy density to replace more conventional materials and to be employed in large scale applications (i.e. electric vehicles) (Mizushima *et al*., 1980). In that respect layered Li-rich metal oxides such as Li_2MnO_2 (David *et al.,* 1983; Johnson *et al.,* 2002; Johnson *et al.,* 2003) are of interest as cathode material for Li ion batteries because they have a high Li concentration, which is a prerequisite for the high energy density.

An efficient way to consider the different candidate materials for energy applications is the use of atomic scale modeling techniques. These can provide information on the defect

Corresponding author: Dr.Navaratnarajah Kuganathan, Department of Materials, Imperial College London, London, SW7 2AZ, United Kingdom and Faculty of Engineering, Environment and Computing, Coventry University, Priory Street, Coventry CV1 5FB, United Kingdom. e-mail:ad0636@coventry.ac.uk

processes including Li ion energetics and mechanisms that can be complementary to experimental work. There is no related theoretical work on the Li related defect process in Li_2MnO_2. The present investigation employs established atomistic modeling techniques based on classical pair potentials to study the intrinsic defect processes, lithium ion diffusion mechanisms and the impact of trivalent doping (Al, Sc, In, Y, Gd and La) in Li_2MnO_2.

2. Computational Methods

All calculations were carried out using the classical pair potential method as implemented in the GULP package (Gale & Rohl, 2003). This method is based on the classical Born model description of an ionic crystal lattice. All systems were treated as crystalline solids with interactions between ions consisting of long-range attractions and short-range repulsive forces representing electron-electron repulsion and van der Waals interactions. The short range interactions were modelled using Buckingham potentials available in the literature (Table 1). Simulation boxes and the corresponding atom positions were relaxed using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm (Gale, 1997). The Mott-Littleton method (Mott & Littleton, 1938) was used to investigate the lattice relaxation about point defects and the migrating ions. It divides the crystal lattice into two concentric spherical regions, where the ions within the inner spherical region (on the order of >700 ions) immediately surrounding the defect relaxed explicitly. Li ion diffusion was calculated considering two adjacent vacancy sites as initial and final configurations. Seven interstitial Li ions were considered in a direct linear route and they were fixed while all other ions were free to relax. The local maximum energy along this diffusion path is calculated and reported as activation energy.

Table 1: *Buckingham interatomic potential parameters {Two-body [Φij (rij) = Aij exp (−rij/ρij) −Cij / rij6]} used in the atomistic simulations of Li_2MnO_2 (Kuganathan & Islam, 2009; Islam et al., 2005; Minervini et al., 1999).*

Interaction	*A* (eV)	*ρ*(Å)	*C* (eV·Å^6)	*Y* (e)	K (eV·Å^{-2})
Li^+ - O^{2-}	632.1018	0.2906	0.000	1.000	99999
Mn^{2+} - O^{2-}	2601.394	0.2780	0.000	3.420	95.0
O^{2-} - O^{2-}	22764.30	0.1490	27.89	−2.860	74.92
Al^{3+} - O^{2-}	1725.20	0.28971	0.000	3.000	99999
Sc^{3+} - O^{2-}	1575.85	0.3211	0.000	3.000	99999
In^{3+} - O^{2-}	1495.65	0.3327	4.33	3.000	99999
Y^{3+} - O^{2-}	1766.40	0.33849	19.43	3.000	99999
Gd^{3+} - O^{2-}	1885.75	0.3399	20.34	3.000	99999
La^{3+} - O^{2-}	2088.79	0.3460	23.25	3.000	99999

3. Results and Discussion

3.1. Structural Modelling

The starting point of this study was to test the quality of the interatomic potentials used in this study by performing a total energy calculation of the Li_2MnO_2 crystal. The layered structure of Li_2MnO_2 (David *et al.,* 1983) exhibits a hexagonal type lattice (space group: $P\bar{3}m1$) as shown in ***Figure 1*** .

Energy minimization calculations were performed on Li_2MnO_2 structure to obtain the equilibrium lattice constants. The calculated equilibrium lattice constants (refer Table 2) are in good agreement with experiment.

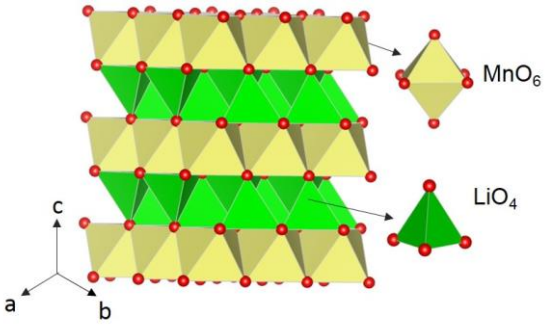

***Figure 1**: Crystal structure of layered Li_2MnO_2 (David et al., 1983). Manganese and lithium atoms form octahedron and tetrahedron respectively with adjacent oxygen atoms*

Table 2: *Calculated and experimental (David et al., 1983) structural parameters for layered hexagonal ($P\bar{3}m1$) Li_2MnO_2*

Parameter	Calculated	Experimental	|Δ|(%)
a (Å)	3.1950	3.1571	1.18
b (Å)	3.1950	3.1571	1.18
c (Å)	5.3030	5.1964	2.01
α = β	90.00	90.00	0.00
γ (°)	120.0	120.0	0.00

3.2. Intrinsic Defects

To calculate the formation energies for Frenkel and Schottky-type defects in Li_2MnO_2, a series of isolated point defect (vacancy and interstitial) energies were calculated. The equations represent the reactions involving these defects as written using Kröger-Vink notation (Kröger & Vink, 1956) and corresponding reaction energies are tabulated in Table 3.

It is calculated that the most favorable intrinsic disorder is the Li-Mn anti-site defect (equation 8). In this defect, Li^+ and Mn^{2+} ions interchange their positions. The exact concentration of this defect is dependent on the temperature and synthesis routes. The formation of Li Frenkel is the second most favorable defect process in this material. As Mn Frenkel, O Frenkel and Schottky defects are highly energetically unfavorable and thus it is unlikely to occur in any significant concentration in Li_2MnO_2.

Table 3: *Energetics of intrinsic defects in Li_2MnO_2.*

Defect Process	Equation	Defect energy (eV)	Defect energy (eV)/defect
Li Frenkel/1	$Li_{Li}^{X} \rightarrow V_{Li}' + Li_i^{\bullet}$	2.59	1.30
O Frenkel/2	$O_O^X \rightarrow V_O^{\bullet\bullet} + O_i''$	9.50	4.75
Mn Frenkel/3	$Mn_{Mn}^{X} \rightarrow V_{Mn}'' + Mn_i^{\bullet\bullet}$	5.29	2.65
Schottky/4	$2\,Li_{Li}^{X} + Mn_{Mn}^{X} + 2\,O_O^X \rightarrow 2\,V_{Li}' + V_{Mn}'' + 2\,V_O^{\bullet\bullet} + Li_2MnO_2$	12.97	2.59
Li_2O Schottky- like/5	$2\,Li_{Li}^{X} + O_O^X \rightarrow 2V_{Li}' + V_O^{\bullet\bullet} + Li_2O$	5.84	1.94
MnO Schottky- like/6	$Mn_{Mn}^{X} + O_O^X \rightarrow V_{Mn}'' + V_O^{\bullet\bullet} + MnO$	7.02	3.51
Li/Mn anti-site (isolated)/7	$Li_{Li}^{X} + Mn_{Mn}^{X} \rightarrow Li_{Mn}' + Mn_{Li}^{\bullet}$	2.90	1.45
Li/Mn anti-site (cluster)/8	$Li_{Li}^{X} + Mn_{Mn}^{X} \rightarrow \{Li_{Mn}' : Mn_{Li}^{\bullet}\}^{X}$	2.00	1.00

3.3. Li ion Diffusion

The intrinsic lithium ion diffusion in and out of the Li_2MnO_2 material determines its use as a possible high-rate cathode material in lithium batteries. The diffusion paths in the Li_2MnO_2 material have not been established experimentally. Using atomistic simulation techniques it is possible to examine various possible diffusion paths responsible for lithium ion diffusion. Two main long range diffusion channels (A and B) connecting local Li hops (see Figure 2(a)) have been identified. The lowest activation energy is calculated to be 0.16 eV for the channel A. Channel A forms a zig-zag pattern with equal Li-Li separation of 2.24 Å. Channel B is slightly different and it exhibits a straight line path with the higher activation energy. Individual Li-Li separations and corresponding activation energy barriers are tabulated in Table 4. Potential energy profile diagram for the Li hopping in channel A is shown in Figure 2(b).

Table 4: *Calculated Li-Li separations and activation energies for the lithium ion migration between two adjacent Li sites as shown in Figure 2a*

Migration path	Li-Li separation (Å)	Activation energy (E_a)(eV)
A	2.24	0.16
B	3.16	2.38

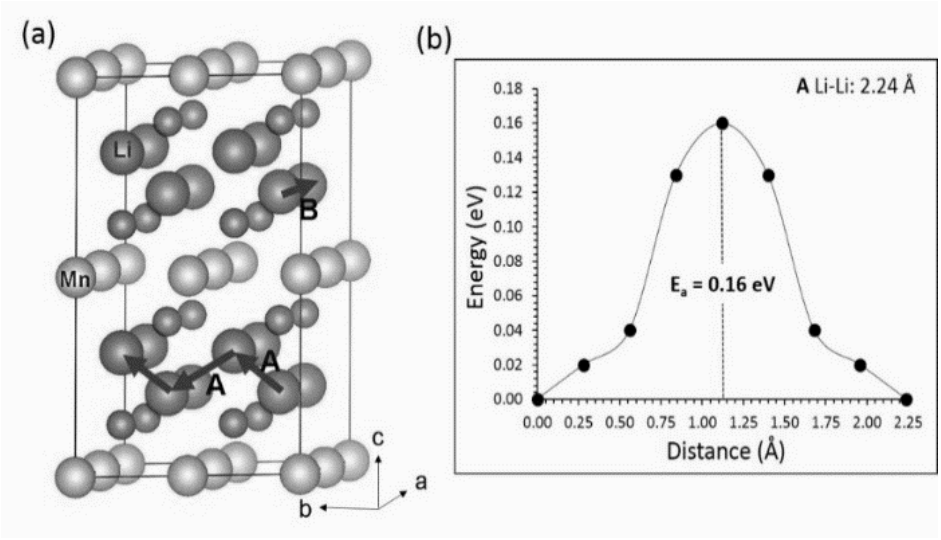

Figure 2: *(a) Lithium vacancy migration paths considered*
(b) Energy profile diagram of Li vacancy hopping between two adjacent Li sites in channel A

3.4. Trivalent Doping

In previous sections, we have shown that Li ions can diffuse with lower activation energy of 0.16 eV. However, the concentration of Li ions in the lattice is dependent on the Li Frenkel energy (1.30 eV/defect) which limits the concentration of V'_{Li}. Here we suggest a way to increase the concentration by substituting trivalent dopants (A) on Mn site (A=Al, Sc, In, Y, Gd and La). This processes can be described as (using the Kröger-Vink notation):

$$A_2O_3 + 2Mn^X_{Mn} + 2Li^X_{Li} \rightarrow 2\,A^{\bullet}_{Mn} + 2\,V'_{Li} + 2MnO + Li_2O \qquad (1)$$

Calculated solution energies together with the ionic radii of trivalent dopants are shown in Figure 3. The lowest solution energy is calculated for Sc^{3+}. This means that the Sc^{3+} is the candidate trivalent dopant that can enhance the concentration of Li vacancy defects in Li_2MnO_2. The second lowest solution energy is calculated for In^{3+} with the energy difference of 0.40 eV compared to that of Sc^{3+}. Other dopants exhibit high positive solution energies meaning that they are less likely to enhance Li vacancy concentration at low temperature.

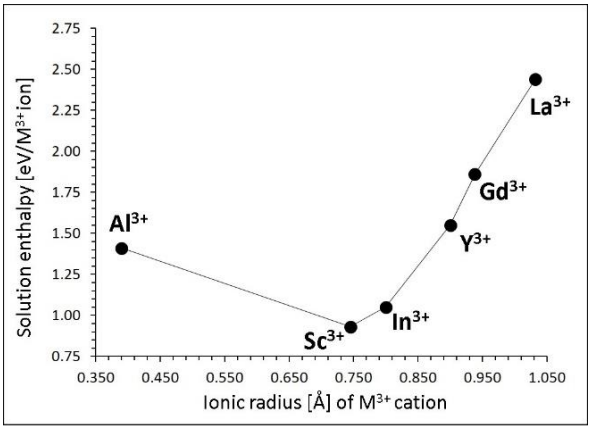

Figure 3: *Enthalpy of solution of A_2O_3 (A = Al, Sc, In, Y, Gd and La) in Li_2MnO_2*

4. Conclusions

Atomistic modelling techniques have been employed to investigate the key issues related to point defects, lithium ion diffusion and dopants in the Li_2MnO_2 material. Our simulation shows good reproduction of the experimental crystal structure of Li_2MnO_2. The most favourable intrinsic disorder type is the Li-Mn anti-site defect. A long range zig-zag Li ion diffusion path with a very low activation energy (0.16 eV) was found. This clearly shows that Li can easily diffuse once the Li vacancies are available. In order to increase the Li vacancies, trivalent dopants were considered. The most energetically favourable dopant on the Mn site is Sc^{3+}.

5. Acknowledgement

Computational facilities and support were provided by High Performance Computing Centre at Imperial College London and the Department of Chemistry, University of Jaffna. Prof. Julian Gale (Faculty of Science and Engineering, Curtain University, Australia) is acknowledged for allowing the members of the department to download and run GULP code on their personal computers.

References

Armstrong, A.R., Kuganathan, N., Islam, M.S. & Bruce, P.G. (2011). Structure and Structure and Lithium Transport Pathways in Li_2FeSiO_4 Cathodes for Lithium Batteries. *Journal of the American Chemical Society, 133*(33), 13031-13035. Retrieved from https://doi.org/10.1021/ja2018543

Bruce, P.G., Freunberger, S.A., Hardwick, L.J. & Tarascon, J.M. (2011). Li–O_2 and Li–S Batteries with High Energy Storage. *Nature Materials, 11*(1), 19-29. doi:10.1038/nmat3191

David, W. I. F., Goodenough, J. B., Thackeray, M. M. & Thomas, M. G. S. R. . (1983). The Crystal structure of Li_2MnO_2. *Revue de Chimie Minerale, 20*(4-5), 636-642.

Fisher, C.A.J., Kuganathan, N. & Islam, M.S. (2013). Defect chemistry and lithium-ion migration in polymorphs of the cathode material Li_2MnSiO_4. *Journal of Materials Chemistry A,, 1*(13), 4207-4214. doi:10.1039/C3TA00111C

Gale, J.D. & Rohl, A.L. (2003). The General Utility Lattice Program (GULP). *Molecular Simulation, 29*(5), 291-341. doi:https://doi.org/10.1080/08927020 31000104887

Gale, J.D. (1997). GULP: A computer program for the symmetry-adapted simulation of solids. *Journal of the Chemical Society, Faraday Transactions, 93*(4), 629-637.

He, X., Zhu, Y. & Mo, Y. (2017). Origin of fast ion diffusion in super-ionic conductors. *Nature Communications, 8*, 15893.

Islam, M.S. & Fisher, C.A.J. (2014). Lithium and sodium battery cathode materials: computational insights into voltage, diffusion and nanostructural properties. *Chemical Society Reviews, 43* (1), 185-204. doi:10.1039/C3CS60199D

Islam, M.S., Driscoll, D.J., Fisher, C.A.J. & Slater, P.R. (2005). Atomic-Scale Investigation of Defects, Dopants, and Lithium Transport in the $LiFePO_4$ Olivine-Type Battery Material. *Chemistry of Materials, 17*(20), 5085-5092. doi:10.1021/cm050999v

Jay, E.E., Rushton, M.J.D., Chroneos, A., Grimes, R.W & Kilner, J.A. (2015). Genetics of superionic conductivity in lithium lanthanum titanates. *Physical Chemistry Chemical Physics, 17*(1), 178-183.

Johnson, C.S., Kim, J-S., Kropf, A.J., Kahaian, A.J., Vaughey, J.T. & Thackeray, M.M. (2002). The role of Li_2MO_2 structures (M=metal ion) in the electrochemistry of $(x)Li_2Mn_{0.5}Ni_{0.5}O_2(1-x)Li_2TiO_3$ electrodes for lithium-ion batteries. *Electrochemistry Communications, 4*(6), 492-498. Retrieved from https://doi.org/10.1016/S1388-2481(02)00346-6

Johnson, C.S., Kim, J-S., Kropf, A.J., Kahaian, A.J., Vaughey, J.T., Fransson, L. M. L., Edström, K. & Thackeray, M.M. (2003). Structural Characterization of Layered $Li_xNi_{0.5}Mn_{0.5}O_2$ ($0 < x \leq 2$) Oxide Electrodes for Li Batteries. *Chemistry of Materials, 15*(12), 2313-2322. Retrieved from https://doi.org/10.1021/cm0204728

Kordatos, A., Kuganathan, N., Kelaidis, N., Iyngaran, P. & Chroneos, A. (2018). Defects and lithium migration in Li_2CuO_2. *Scientific Reports, 8*(1), 6754. doi:10.1038/s41598-018-25239-5

Kröger, F.A. & Vink, H.J. (1956). Relations between the Concentrations of Imperfections in Crystalline Solids. In F. Seitz, & D. Turnbull (Eds.), *Solid State Physics* (Vol. 3, pp. 307-435). Massachusetts: Academic Press. Retrieved from https://doi.org/10.1016/S0081-1947(08)60135-6

Kuganathan, N. & Chroneos, A. (2018a). Defects, Dopants and Sodium Mobility in Na_2MnSiO_4. *Scientific Reports, 8*, 14669. doi:10.1038/s41598-018-32856-7

Kuganathan, N. & Islam, M.S. (2009). Li_2MnSiO_4 Lithium Battery Material: Atomic-Scale Study of Defects, Lithium Mobility, and Trivalent Dopants. *Chemistry of Materials, 21*(21), 5196-5202. Retrieved from https://doi.org/10.1021/cm902163k

Kuganathan, N., Ganeshalingam, S. & Chroneos, A. (2018b). Defects, Dopants and Lithium Mobility in $Li_9V_3(P_2O_7)3(PO_4)_2$. *Scientific Reports, 8*, 8140. doi: 10.1038/s41598-018-26597-w

Kuganathan, N., Iyngaran, P. & Chroneos, A. (2018c). Lithium diffusion in Li_5FeO_4. *Scientific Reports, 8*, 5832. doi:10.1038/s41598-018-24168-7

Kuganathan, N., Kordatos, A. & Chroneos, A. (2018d). Li_2SnO_3 as a Cathode Material for Lithium-ion Batteries: Defects, Lithium Ion Diffusion and Dopants. *Scientific Reports, 8*, 12621.

Kuganathan, N., Kordatos, A., Fitzpatrick, M.E., Vovk, R.V. & Chroneos, A. (2018e). Defect process and lithium diffusion in Li_2TiO_3. *Solid State Ionics, 327*, 93-98. doi:https://doi.org/10.1016/j.ssi.2018.10.030

Minervini, L. Z. (1999). Defect cluster formation in M_2O_3-doped CeO_2. *Solid State Ionics, 116*(3-4), 339-349.

Mizushima, K. J. (1980). LI_XCOO_2 $(0<X<-1)$: A new cathode material for baterries of high energy density. *Materials Research Bulletin, 15*(6), 783-789.

Mott, N. F. (1938). Conduction in polar crystals. I. Electrolytic conduction in solid salts. *Transactions of the Faraday Society, 34*, 485-499. doi:10.1039/TF9383400485

Tarascon, J. &. (2001). Issues and challenges facing rechargeable lithium batteries. *Nature, 414*, 359–367.

Research Article

Water Quality Assessment of the Pond "Ariyakulam", Jaffna - A view through Correlation Study and Regression Analysis

Arthiya Rajasingham, Abyerami Sivaruban*

Department of Zoology, University of Jaffna, Sri Lanka.

Received: 24 April, 2018, Accepted: 19 March, 2019

Abstract:

The present study was designed to assess the physico-chemical status of the pond *"Ariyakulam"*, to compare its status with surface water quality standards and quantify the impacts of these parameters on each other though the regression and correlation coefficient methods. Eleven selected parameters such as Air Temperature (AT), Water Temperatures (WT), Dissolved Oxygen (DO), Total Dissolved Solid (TDS), Electrical Conductivity (EC), Salinity, pH, Chemical Oxygen Demand (COD), Biological Oxygen Demand (BOD), Nitrate nitrogen (NO_3^-), Ammonia nitrogen (NH_4^+) were taken for the analyses. Though the pond covers approximately an area of 12,000 m^2 located in a historical place, proper maintenance is lacking and the preservation of the beauty of this water body is largely neglected. Water samples were collected twice a month for a period of 8 months from August 2016 to January 2017 and then January to February, 2018 from 0700 to 0800 hrs in eight locations within the pond. The values of TDS, DO, EC, and NH_4^+ indicated that the pond "*Ariyakulam*" was in polluted state. The correlation coefficient values and regression analysis methods revealed the relationship among the parameters. The results from these methods demonstrate a user friendly measure for a rapid water quality assessment and continuous monitoring programmes.

Keywords: Water quality, Surface water, Physico-chemical parameters, Correlation, Regression

1. Introduction

Inland waters of Sri Lanka could be categorized as surface and ground water. On an average Sri Lanka internally produced 52 billion cubic meters surface water out of a total of 52.8 billion cubic meters renewable water resources per year (RWR) (Food and Agriculture Organization, 2016). As far as Sri Lanka is concerned, 4.43% (2905 sq Km) of the land area is covered by the surface water (Ministry of Forestry and Environment, Sri Lanka, 2001). Out of which Jaffna Peninsula including islands covering a land area of 1200 Km^2 possesses a total of 297 stagnant water bodies (Piyasiri, 2009). Ponds, rivers and lagoons are the sources of surface waters. Because surface waters significantly contribute to the replenishing of ground waters, maintenance of the surface water quality standards is of utmost importance to the healthy environment.

The pond "Ariyakulam" ($9^0 40' 5.8'' N$ and $80^0 1' 8.4'' E$) could be assigned to the Class III category according to the classification of surface waters (United States Environmental

Corresponding author: Dr.Abyerami Sivaruban, Department of Zoology, University of Jaffna, P.O.Box 57, Jaffna, Sri Lanka.
e-mail: abyerami@gmail.com

Protection Agency, 2014). After the incidence of massive fish kills which occurred in March, 2016, a study was designed to find out the physico-chemical status of the pond and also to compare its status with Sri Lankan and international surface water quality standards.

"*Ariyakulam*" is located close to the Jaffna town at the Stanley road, Point Pedro road junction and covers approximately 12,000 m^2 in area. It is one of the oldest ponds within the water drainage systems that are in operation in Jaffna. It receives water mainly from the seasonal rains and the excess of water from the nearby ponds through the connecting canal. In the same way overflowing water from the "*Ariyakulam"* runs into the other pond finally reaching the Jaffna lagoon at Pannai. Further crowded human settlements are seen along the Eastern and Northern boarders and Western and Southern sides facing the roads.

Inspite of the historical significance of the pond, anthropogenic impacts were heavy in the sense that *"Ariyakulam*" receives daily influx of waste waters from nearby settlements and plastics and other wastes seen inside the pond. Its surface was covered with aquatic plants such as lotus, *Azolla* and *Lemna* species. Even though the pond is not clean, it harbours number of organisms such as fishes, water snakes, avifauna, invertebrates and planktonic organisms.

Though few research papers were published with regard to "*Ariyakulam"* (Balasundaram, 1986; Selvarajah & Rajaratnam, 1990; Patrick *et al*., 2012; Kadotgasan, 2010; Abyerami & Arthiyan, 2016), no detailed report on the assessment of the water quality of the pond has been done so far. As "*Ariyakulam*" is situated around a high water usage area, proper maintenance of water quality of this pond is very much essential. As such a study was designed to assess the water quality so that the proper planning can be formulated to maintain this water body. The objectives of the study were assessment of the present water quality through the analyses of the selected water quality parameters; determination of their relationships and finally comparing them with the surface water quality standards.

2. Materials and Methods

2.1 Collection of water samples:

Water samples were collected twice a month for a period of 8 months from August 2016 to January 2017 and then January to February 2018, from the pond "*Ariyakulam*". Clean buckets were dipped 1m away from its boundary and samplings were done from 0700 to 0800 hrs in eight locations namely L1 - L8 in order to cover entire pond **(*Figure 1*)**.

Figure 1 **:** *Satellite image of the pond "Ariyakulam"*

Eleven physico-chemical parameters were measured using standard protocols (**Table 1**). Dissolved Oxygen (DO), Air Temperature (AT) and Water Temperatures (WT) were measured on the site. Samples were brought to the laboratory and kept in the ambient temperature until processing the rest of the parameters such as Total Dissolved Solids (TDS), Electrical Conductivity (EC), Salinity, pH, Chemical Oxygen Demand (COD), Biological Oxygen Demand (BOD), Nitrate Nitrogen (NO_3^-), and Ammonia Nitrogen (NH_4^+). All analyses except BOD were completed within 24 hours of collection. (**Table 2)**

Table 1: *Physico-chemical parameters and their standard protocols*

Parameter	Method
Dissolved Oxygen (DO)	HQ30d probe
Air Temperature (AT)	Centigrade thermometer
Water Temperatures (WT)	Centigrade thermometer
Total Dissolved Solid (TDS) Electrical Conductivity (EC) Salinity	HQ40d probe
pH	Sension+ pH probe
Chemical Oxygen Demand (COD)	APHA, 1999
Biological Oxygen Demand (BOD)	HQ30d probe after three days of sample collection
Nitrate nitrogen (NO_3) Ammonia nitrogen (NH_4^+)	Kjeldhal distillation method

2.2 Statistical analysis

The regression analysis was done to find out the linear relationship among the parameters by determining the correlation coefficient **Table 3**. Pearson correlation coefficient was calculated using computer software MS Excel 2010, and a correlation matix was obtained **Table 4.** This analysis attempts to establish the nature of the relationship between the variables and thereby providing a mechanism for prediction or forecasting the status of the water resource (Kumar & Sinha, 2010). A handful of papers (Daraigan *et al.*, 2011; Meena & Bhargava, 2012; Heydari *et al.*, 2013; Chaubey & Patil, 2015) assessed the drinking water quality by this method. This paper utilizes the same methods to assess the surface water quality so that planning of rapid water quality management programs can be done easily.

3. Results and Discussion

The summary of monthly variation of the physico-chemical parameters of the pond (with standard error) and Sri Lankan and International surface water quality standards were given in the **Table 2.** The statistical differences were calculated for these and given in the **Table 5**.

3.1 Physico-chemical parameters

Air temperature showed 5^{o}C fluctuation from 28.5 (± 1.13) to 31^{0}C (± 1.07) and a little lesser fluctuation (< than 3^{o}C) from 27.69 (± 1.13) to 28.56 (± 1.13) was seen in the water temperature. Both parameters showed statistically significant variations among the months.

Dissolved oxygen significantly varied from 1.06 (± 0.41) - 5.08 mg/l (± 1.13). Only during the pre monsoon and monsoon period, the DO of the pond was reaching the Sri Lankan standards (4mg /l). ie the value was above the standards during October 2016 (5.08 mg/l ± 1.13) and closed to the standards during December 2016 (3.60 mg/l ± 1.08) and January 2017 (3.76 mg/l ± 0.67).

Statistically significant differences in pH were observed and it showed the basic range {8.2 (± 0.3) - 9.4 (±0.2)} throughout the study period. Kannaiyan *et al.,* (1983) reported that the high pH range (> 8.0) resulted because of the growth and reproduction of algae when they use CO_2 and this reduction causes the pH to increase. The pond showed the phenomenal growth of *Azolla, Lemna* and Lotus and this might be the reason for the higher pH values obtained during the study period.

Table 2: *Average results of physico-chemical parameters of the pond "Ariyakulam" from August 2016 - January 2017 and January - February 2018*

Year and Month	AT °C	WT °C	pH	EC µS/cm	Values in mg / l				Salinity	Values in mg / l.	
					TDS	DO	BOD_3	COD		NO_3^-	$NH4^+$
Surface water quality standards for Class III	-	<40 (4)	6.0 - 9.0 (1) 6.5 - 8.5 (3)	1275 (2)	9.67 (2) 1300 (3)	4 (1) ≥ 5 (3) 5 (4)	5 (1) 6 (3)	NA (2) 15 (3)	-	10 (1) 5.6 (3)	≤ 0.02 (2) 0.8 - NH_4 (3)
August 2016	28.5 ± 1.13	27.69 ± 1.13	8.2 ±0.3	1336.88 ± 167.63	685.88 ± 83.96	1.06 ± 0.41	0.40 ± 0.34	3.5 ±0.71	0.62 ± 0.04	0.73	1.52
September 2016	29 ± 1.36	28.50 ± 1.36	9.2 ± 0.2	1056.75 ± 79.78	517.88 ± 41.05	1.38 ± 0.88	0.80 ±1.04	3.5 ± 0.58	0.62 ± 0.04	0.73	1.62
October 2016	30 ± 1.02	28.50 ± 1.02	8.7 ± 0.1	1515.19 ± 93.57	742.44 ± 39.85	5.08 ± 1.13	0.80 ±1.04	1.3 ± 0.72	0.62 ± 0.06	0.82	1.43
November 2016	31 ± 1.07	28.56 ± 1.07	8.8 ± 0.2	1415.50 ± 162.93	616.06 ± 135.41	2.71 ± 1.04	1.75 ± 0.73	1.8 ± 0.79	0.64 ± 0.06	1.21	1.82
December 2016	29.5 ± 0.94	28.19 ± 0.94	8.4 ± 0.4	1326.44 ± 93.79	529.13 ± 89.64	3.60 ± 1.08	1.92 ± 0.31	1.3 ± 0.59	0.63 ± 0.05	0.98	1.75
January 2017	30 ± 0.98	28.31 ± .98	8.7± 0.2	1347.69 ± 01.00	801.44 ± 24.50	3.76 ± 0.67	2.01 ± 0.61	2.8 ± 0.83	0.76 ± 0.06	1.02	1.84
January 2018	31 ± 1.07	27.69 ± 1.13	9.2 ± 0.2	742.44 ± 39.85	1515.49 ± 93.77	0.84 ± 1.14	3.5 ± 0.58	5.28 ± 1.43	0.64 ± 0.06	1.31	1.86
February 2018	30 ± 1.02	28.56 ± 1.07	9.4 ± 0.2	732.34 ± 38.85	1615.19 ± 92.57	0.82 ±1.01	3.4 ± 0.43	5.20 ± 1.18	0.64 ± 0.06	1.54	1.75

(1) Sri Lanka Standards Institution., 1985 ;
(2) United States Environmental Protection Agency,2014;
(3) Buijs & Toader, 2007;
(4) Reebs, 2009

Table 3: *Linear correlation coefficient 'r' and regression equation for some pairs of parameters with significant value of correlation ($p \leq 0.1$)*

Parameter	r value	Regression coefficient		Regression equation	F value	p value
		A	b			
AT- DO	0.653	-26.5028	0.9921	AT = -26.5028 DO + 0.9921	4.480	0.078
AT- EC	0.572	-1424.841	92.9634	AT= -1424.841 EC + 92.9634	2.919	0.138
TDS – DO	0.5619	-0.8901	0.0058	TDS = -0.8901 DO+ 0.0058	2.769	0.147
AT- NH_4^+	0.620	1.0163	0.7070	AT= .0163 NH_4^+ + 0.7070	3.75	0.100
pH - NO_3^-	0.607	0.7689	0.0168	pH= 0.7689 NO_3^-+ 0.0168	3.517	0.109
DO – EC	0.802	1132.1848	68.5242	DO= 1132.1848 EC+ 68.5242	10.839	0.016
DO - NO_3^-	0.656	0.7912	0.0422	DO =0.7912 NO_3^-+0.0422	4.548	0.076
BOD – SALINITY	0.606	0.5853	0.0491	BOD =0.5853 SALINITY +0.0491	3.486	0.111
BOD - NH_4^+	0.612	1.3874	0.2155	BOD =1.3874 NH_4^++ 0.2155	3.594	0.106
EC - NO_3^-	0.712	0.2457	0.0005x	EC =0.2457 NO_3^-+ 0.0005	6.196	0.047

Table 4 :Pearson Correlation (r) matrix for the abiotic factors

Parameter	AT	WT	p^H	TDS	DO	BOD	COD	EC	Salinity	$NO3^-$	NH4
AT	1										
WT	0.163	1									
Ph	0.440	0.363	1								
TDS	0.337	-0.185	0.103	1							
DO	0.653	0.121	0.433	0.561	1						
BOD	0.366	0.316	-0.206	-0.119	0.100	1					
COD	-0.236	-0.391	0.436	0.202	-0.223	-0.545	1				
EC	0.572	-0.030	0.203	0.669	0.802	-0.095	-0.136	1			
Salinity	0.212	0.074	-0.023	0.532	0.162	0.606	0.075	0.012	1		
NO_3^-	0.696	0.132	0.607	0.341	0.656	0.163	0.149	0.712	0.155	1	
NH_4^+	0.620	-0.044	0.352	0.079	0.267	0.612	0.150	0.146	0.517	0.687	1
Strong 1; Moderate 14; Weak 28 ; Negative 12											

Table 5: *Statistical analysis of physico-chemical parameters*

Parameter	Max	Min	Mean	Range	p value ($p \leq 0.1$)
AT	31.00	28.50	29.88	2.50	6.060e-09
WT	28.56	27.69	28.25	0.87	8.949e-09
pH	9.40	8.20	8.83	1.20	2.414e-05
TDS	801.44	517.88	670.95	283.56	0.00
DO	5.28	1.06	3.51	4.22	0.004
BOD	2.01	0.40	1.17	1.61	0.140
COD	3.50	1.30	2.64	2.20	0.016
EC	1615.19	1391.14	1391.14	224.05	0.00
Salinity	0.76	0.62	0.65	0.14	0.268
(NO_3^-)	1.54	0.73	1.04	0.81	0.166
(NH_4^+)	31.00	1.43	1.70	29.57	0.066

The concentrations of nitrate nitrogen was within the limits (0.73mg/l - 1.54mg/l), while the ammonia nitrogen was exceeded the standards (1.43mg/l - 1.86mg/l). The maximum permissible level of nitrate nitrogen and ammonia nitrogen for the surface water are 10 mg/l and 0.02 - 0.8 mg/l respectively. In the environment, ammonia generates from the agricultural and industrial processes. But this pond has possible pollution of fecal contaminations, sewage and animal wastes by the nearby human settlements over the others and fecal contamination was previously recorded from the same pond (Abyerami & Arthiyan, 2016). High levels of ammonia with corresponding low level of DO indicate the prevailing decomposition process in the pond.

EC values showed fluctuation from 732.34 µS/cm (±38.85) - 1515.19 µS/cm (± 93.57) and varied significantly among months ($p < 0.1$). The standard EC value reported was 1275 µS/cm (United States Environmental Protection Agency, 2014). However the conductivity of most freshwaters ranges from 10 to 1,000 µS cm-1 but may exceed 1,000 µS/cm especially in polluted water bodies or those receiving large quantities of land run-off (Meybeck & Helmer, 1996). Further these values indicate high mineralization state of the surface water area and this may be due to high concentration of ionic constituents present in the water bodies (Salam *et al*., 2012).

TDS value of the pond was ranged between 517.88 mg/l (± 41.05) to 1615.19 mg/l (± 92.57). During January and February 2018, TDS was higher than the permissible limit of 1300mg/l. This indicates the turbid nature of the pond which obviously limits the level of light penetration into the pond.

BOD ranged from 0.40 mg/l (± 0.34) - 3.5 mg/l (± 0.58) which was found within the standards. For most instances tolerable limits of BOD ranges from 5 - 6 mg/l. However if the BOD value exceeds 4mg/l the water body can be considered as polluted (Piyasiri, 2009). COD varied from of 1.31mg/l (±0.59) to 5.28mg/l (±1.43) which also found to be within the standards.

3.2 Correlation matrix and regression model

The numerical values of correlation coefficient (r) for the eleven water quality parameters are tabulated in **Table 2** Correlation analysis measures the closeness of the relationship between chosen independent and dependent variables. If the correlation coefficient is nearer to (+1) to (-1), then it shows the probability of the linear relationship. The greater the value of regression coefficient, the better is the fit and more useful the regression variables (Kumar & Sinha, 2010).

To find about the relationship among these parameters, the regression equation was used where the correlation coefficient (r) measures the degree of association that exist between any two variables. The correlation considered as strong when it is in the range of +0.8 to 1.0 and -0.8 to -1.0; moderate when its value falls between 0.5 and 0.8 and -0.5 to -0.8; when the range from 0.5 to +0.0 and 0.0 to -0.5, then it is week. Actually correlation is the mutual relationship between two variables. Direct correlation exists when increase or decrease in the value of one parameter is associated with a corresponding increase or decrease in the value of other parameter (Patil & Patil, 2011).

It is shown that a high positive (strong) correlation existed between DO and EC (R=0.802). $N0_3^-$, DO and p^H were positively correlated with the all the water quality parameters.

The negative correlations were found in 12 cases between AT and COD, WT and COD, WT and TDS, WT and EC, WT and NH_4^+, P^H and BOD, P^H and Salinity, TDS and BOD, DO and COD, BOD and COD, BOD and EC and COD and EC. Poor positive correlation (moderate) was found in 14 cases. Very negligible positive (weak) correlation was observed in 28 cases where as the r value was less than 0.5

In the current study it is evident that AT, P^H, DO, EC were strongly correlated with NO_3^- ($r > 0.6$). NH_4^+ was significantly correlated with AT, NO_3^-, BOD and salinity ($r > 0.6$). Further TDS was significantly correlated with DO, EC and salinity ($r > 0.5$). The following two regression relationship have more or less same correlation coefficient value such as AT with BOD ($r=0.366$) and WT with P^H ($r=0.363$). Likewise P^H and Nitrate content ($r = 0.607$), BOD with salinity ($r = 0.606$) **(Table 4**).

In 10 cases correlation coefficient for the water quality parameters found to have better and higher level of significance ($p \leq 0.1$) were shown in **Table 3.** In the current study, it is evident that the AT, DO, EC, TDS, $NH4^+$, $NO3^-$, BOD and salinity were significantly correlated ($r>0.6$) among themselves.

4. Conclusions

The values of TDS, DO, EC, and NH_4^+ indicates that the pond "*Ariyakulam*" is in polluted state with high level of ionic compounds. COD values obtained in January and February 2018 also agreed with this. Lower level of DO in most instances suggested survival risks of pond organisms during the hotter months. The linear correlation is very useful to get fairly accurate idea of quality of water by determining a few parameters experimentally.

It can be concluded that the AT, TDS, DO, EC and nitrogen contents are important physico-chemical surface water quality parameters, because they are correlated with most of the water parameters.This work could be more enhanced by studying ground water quality of the adjacent areas. Emphasis should be given to reduce the P^H, EC, TDS, NH_4^+ levels and whenever COD crosses the permissible level.

Finally if all or parts of these are implemented, the pond "Ariyakulam" can be converted as a very good recreational area for the public at the same time a healthy environment can also be achieved.

References

Abyerami, S. & Arthiyan, S. (2016). Occurrence of fish kills in "Ariyakulam" during early March, 2016 and measures for the prevention of reiterate phenomena. *Proceedings of the Jaffna Science Association. 23*, p. 7. Jaffna: Jaffna Science Association.

Balasundaram, C. (1986). *The distribution of copepod in some ponds in Jaffna Peninsula.* B.Sc. dissertation, University of Jaffna.

Buijs, P. & Toader, C. (2007). *Proposed system of surface water quality standards for Moldova: Technical report.* Organization for Economic Co-operation and Developmemt. Retrieved from http://www.oecd.org/environment/outreach /38120922.pdf

Chaubey, S. & Patil, M.K. (2015). Correlation Study and Regression Analysis of Water Quality Assessment of Nagpur City, India. *International Journal of Scientific and Research Publications, 5*(11), 753 – 757.

Daraigan, S.G., Wahdain, A.S., Ba-Mosa, A.S., & Obid, M.H. (2011). Linear correlation analysis study of drinking water quality data for Al - Mukalla City, Hadhramout, Yemen. *International Journal of Environmental Sciences, 1*(7), 1692-1701.

Food and Agriculture Organization. (2016). AQUASTAT Main Database, Food and Agriculture Organization of the United Nations. Retrieved from www.fao.org/nr/water/aquastat/data/query/index.html?lang=en

Heydari, M.M., Abbasi, A., Rohani , S.M. & Hosseini, S.M.A. (2013). Correlation Study and Regression Analysis of Drinking Water Quality in Kashan City, Iran. *Walailak Journal of Science and Technology, 10*(3), 315-324.

Kadotgasan, J.M. (2010). *Study the zooplankton distribution in the Northern tropical ponds in Sri Lanka.* M.Sc. thesis, University of Jaffna.

Kannaiyan,S., Viswanath,G. & Rajeswari, N. (1983). Influence of green and brown colour variation on NPK Contents of Azolla pinnata. *L.Sci.Cult, 49*, 251-252.

Kumar, N. & Sinha, D.K. (2010). Drinking water quality management through correlation studies among various physico-chemical parameters: a case study. *International Journal of Environmental Sciences, 1*(2), 253 - 259.

Meena, B.S. & Bhargava, N. (2012). Physico-chemical characteristics of groundwater of some villages of dag block in Jhalawar district of Rajasthan state (India). *Rasayan Journal of Chemistry, 5*(4), 438-444.

Meybeck, M. & Helmer, R. (1996). An Introduction to Water Quality. In D. Chapman (ed.), *Water Quality Assessments - A Guide to the Use of Biota, Sediments and Water in Environmental Monitoring* (pp. 1-22). London: E & FN Spon.

Ministry of Forestry and Environment, Sri Lanka & Development, Environment and Management Associates, Sri Lanka. (2001). *Srilanka: State of the Environment 2001.* Pathumthani, Thailand: United Nations Environment Programme.

Patil, V.T. & Patil, P.R. (2011). Groundwater quality of open wells and tube wells around Amalner town of Jalgaon district, Maharashtra, India. *Electronic Journal of Chemistry, 8*(1), 53-58.

Patrick, A.E.S., Kadotgasan, J.M. & Naveendrakumar, G. (2012). Study to detect impacts of pollution on the distribution of Zooplankton in the Northern tropical ponds in Sri Lanka. *Archives of Applied Science Research, 4*(6), 2552-2556.

Piyasiri, S. (2009). Surface waters, their status and management. *Economic Review, June/July*, 23-31.

Reebs, S.G. (2009). Oxygen and fish behavior. Retrieved from www.howfishbehave.ca/pdf/oxygen.pdf

Salam, S.M.A., Mollah, M.A., Tasnuva, A. & Zaman M.R. (2012). Physico-chemical Evaluation of Ground and Surface Water of Mohanpur Upazila of Rajshahi District. *Journal of Environmental Sciences and Natural Resources, 5*(2), 275-280.

Selvarajah, N. & Rajaratnam,V. (1990). Limnology of two temporary ponds in the Jaffna Peninsula, Sri Lanka. *Vingnanam Journal of Science., 5*, 7-13.

Sri Lanka Standards Institution. (1985). *SLS: 722: 1985: Tolerance limit for inland surface waters used as raw water for public water supply.* Colombo: Sri Lanka Standards Institution.

United States Environmental Protection Agency. (2014). *Chapter 62- 302: Water quality standards.* Washington, D.C.: United States Environmental Protection Agency. Retrieved from https://www.epa.gov/sites/production/files/2014-12/documents/fl_section62-302.pdf.